

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 30, 2004, 06:15:28 ; Search time 5051 Seconds  
(without alignments)  
3192.593 Million cell updates/sec

Title: US-10-088-187a-11

Perfect score: 1796

Sequence: 1 MPRPFHKLIFSSTIQEKL.....LRTDRFVLKVTAFRNXYV 341

Scoring table: BLOSUM62

-Q-/cgn2 1/USPTO spool/US10088187/runat 29122004 115842 119/app.query.fasta\_1.519

-DB=GenEmbl -QFMT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPEXT=0

-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000

-USER=US10088187 @CNG 1 1 3731 @runat 29122004 115842 119 -NCPU=6 -ICPU=3

-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG

-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh

-Q-/cgn2 1/USPTO spool/US10088187/runat 29122004 115842 119/app.query.fasta\_1.519

-DB=GenEmbl -QFMT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPEXT=0

-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000

-USER=US10088187 @CNG 1 1 3731 @runat 29122004 115842 119 -NCPU=6 -ICPU=3

-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG

-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Scoring table: BLOSUM62

-Q-/cgn2 1/USPTO spool/US10088187/runat 29122004 115842 119/app.query.fasta\_1.519

-DB=GenEmbl -QFMT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPEXT=0

-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000

-USER=US10088187 @CNG 1 1 3731 @runat 29122004 115842 119 -NCPU=6 -ICPU=3

-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG

-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Scoring table: BLOSUM62

5	1768	98.4	1494	6	AX101039	Sequence
6	1614.5	89.9	5000	8	AF289051	Arabidops
7	1614.5	89.9	79186	8	AP000735	Arabidops
8	1601.5	89.2	5000	6	AX101027	Sequence
9	1513	84.2	990	8	AY517929	Brassica
10	1509	84.0	990	8	AY356368	Brassica
11	1113.5	62.0	1146	8	AY042868	Arabidops
12	1098	61.1	125021	8	AC007504	Arabidops
13	946	52.7	681	6	AX505578	Sequence
14	850	47.3	78589	8	AP006145	Lotus cor
15	808	45.0	243777	2	AC137825	Medicago
16	359	20.0	122693	8	AC145449	Medicago
17	357	19.9	114098	8	AC144459	Medicago
18	333	18.5	48368	2	AC140105	Medicago
19	332.5	18.5	92281	8	AP004896	Lotus cor
20	301.5	16.8	48368	2	AC140105	Medicago
21	298.5	16.6	1361	6	AX558824	Sequence
22	293.5	16.3	1041	6	AX558782	Sequence
23	281	15.6	1762	8	AK101941	Oryza sat
24	281	15.6	1884	8	AK067571	Oryza sat
25	270	15.0	1469	8	AK108660	Oryza sat
26	249.5	13.9	129647	8	AC097280	Oryza sat
27	249	13.9	108028	8	AC135233	Medicago
28	239	13.3	1014	8	ATH441117	Arabidops
29	237.5	13.2	129204	8	AC146747	Medicago
30	236.5	13.2	1409	8	AK063506	Oryza sat
31	236	13.1	132741	8	CNS08CAA	Arabidops
32	229.5	12.8	2000	6	AX508278	Sequence
33	229	12.8	573	8	BT010604	Arabidops
34	220	12.2	204120	8	AY661859	Sorghum b
35	217	12.1	661	8	BT008541	Arabidops
36	217	12.1	809	8	BT006139	Arabidops
37	202	11.2	1495	8	AK109729	Oryza sat
38	195.5	10.9	937	8	BT005553	Arabidops
39	195.5	10.9	1113	8	BT004099	Arabidops
40	192.5	10.9	1221	8	AY087853	Arabidops
41	192.5	10.7	90596	8	BT15B16	Arabidops
42	192.5	10.7	198697	8	ATCHRIV4	Arabidops
43	191.5	10.7	140383	8	CNS08CD5	Arabidops
44	188.5	10.5	125021	8	AC007504	Arabidops
45	185.5	10.3	36495	8	AB015473	Arabidops

# ALIGNMENTS

RESULT 1	CQ805406	Sequence 1817 from Patent WO2004035798.	1026 bp	DNA	linear	PAT 10-MAY-2004
LOCUS	CQ805406	Sequence 1817 from Patent WO2004035798.	1026 bp	DNA	linear	PAT 10-MAY-2004
DEFINITION	CQ805406	Sequence 1817 from Patent WO2004035798.	1026 bp	DNA	linear	PAT 10-MAY-2004
ACCESSION	CQ805406	Sequence 1817 from Patent WO2004035798.	1026 bp	DNA	linear	PAT 10-MAY-2004
VERSION	CQ805406.1	GI:47111337	1026 bp	DNA	linear	PAT 10-MAY-2004
KEYWORDS	Arabidopsis thaliana (thale cress)					
SOURCE	Arabidopsis thaliana					
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi					
REFERENCE	Inze, D., de Veylder, L. and Vlieghe, K.					
AUTHORS	Identification of novel e2f target genes and use thereof					
TITLE	Patent: WO 2004035798-A 1817 29-APR-2004;					
JOURNAL	CropDesign N.V. (BE)					
FEATURES	Location/Qualifiers					
source	1..1026					
	/organism="Arabidopsis thaliana"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:3702"					
ORIGIN						
Alignment Scores:						
Pred. No.:	1.71e-139	Length:	1026			
Score:	1796.00	Matches:	341			
Percent Similarity:	100.00%	Conservative:	0			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1796	100.0	1026	CQ805406 Sequence
2	1796	100.0	1495	AX101036 Sequence
3	1796	100.0	1538	AF289052 Arabidops
4	1781	99.2	1495	AX101038 Sequence



Db 569 GAGATCAATACCATTCACCGGTCTCATGGATTCGGCTCACACCACTTCAACACGGCC 628  
Qy 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerValTyr 140  
Db 629 CGTTGTTTGAAGACCTTGAAGATGAAGATGCCGAGGTCACTTCTCTCTCTCTGTAC 688  
Qy 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160  
Db 689 CCATCACCACTTCTCGAGTCTACAGTACAGCAGCAACAAAGGGTATGCTAGTTCAGGCATC 748  
Qy 161 GlnThrLeuPheThrGlyProValLysAlaGluGluProThrProThrProLysIlePro 180  
Db 749 CAAACCTTGTCTACTGACCACTTAAGCTGAAGGCCAACGCCCAACCCCAAAATACCT 808  
Qy 181 LysLeuArgGlyArgLysLysAsnAlaAspProGluGluIleAsnSerAlaPro 200  
Db 809 AAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 868  
Qy 201 ArgAspAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220  
Db 869 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGAGTGTCTCGAGAAAGAGA 928  
Qy 221 ThrValThrAlaGluGluArgAlaIleAsnAlaLysThrPheGluProThr 240  
Db 929 ACCGTGACTGCAGAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 988  
Qy 241 AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260  
Db 989 AACCTTTCTTCAGAGTGTCTCGAGGATCTCTATCATACAGAGTGTCTCATCATAT 1048  
Qy 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyLysSerGlyPheIleLysValGln 280  
Db 1049 CTTCTCTCTGGGTTTCTGAGAGTACCTCAAGTGGGATCTCGGGTTCTCAAGTCCAG 1108  
Qy 281 LeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSer 300  
Db 1109 CTTGCGGAGAAACAAATGGCTGTTCGATGTCTCTACAAAGCCGGAGAGCCAAATTCAGT 1168  
Qy 301 GlnGlyTyrTyrGluPheThrLeuGluAsnAsnLeuGlyGluGlyAspValCysValPhe 320  
Db 1169 CAAAGATGTACGATCTCTAGAGAACCACTTAGGAGAGGAGGAGGAGGAGGAGGAGG 1228  
Qy 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyr 340  
Db 1229 GAGCTGCTCAGAACCAAGATTCGTTTGAAGTGCACAGCCTTTCGAGTCAACGAGTAC 1288  
Qy 341 Val 341  
Db 1289 GTC 1291

RESULT 3  
LOCUS AF289052  
DEFINITION Arabidopsis thaliana reduced vernalization response 1 (VRN1) mRNA,  
VRN1-Ler allele, complete cds.  
ACCESSION AF289052  
VERSION AF289052.1 GI:21734795  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
1 (bases 1 to 1538)  
REFERENCE  
AUTHORS Levy,Y.Y., Mesnage,S., Wylne,J.S., Gendall,A.R. and Dean,C.  
TITLE Multiple roles of Arabidopsis VRN1 in vernalization and flowering  
time control  
JOURNAL Science 297 (5579), 243-246 (2002)  
MEDLINE 22111275  
PUBMED 12114624  
REFERENCE  
AUTHORS Levy,Y.Y., Gendall,A.R. and Dean,C.

VRN1, a gene required for response to vernalization  
Unpublished  
3 (bases 1 to 1538)  
Levy,Y.Y. and Dean,C.  
Direct Submission  
Submitted (23-JUL-2000) Molecular Genetics, John Innes Centre,  
Colney Lane, Norwich, Norfolk NR4 7UH, UK  
FEATURES  
source  
1..1538  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/map="between m1339 and m560B2"  
/ecotype="Landsberg erecta"  
1..1538  
/gene="VRN1"  
/allele="Ler"  
312..1337  
/gene="VRN1"  
/allele="Ler"  
/note="contains two VP1/ABI3-like B3 domains"  
/codon\_start=1  
/product="reduced vernalization response 1"  
/protein\_id="AAW76973.1"  
/db\_xref="GI:21734796"  
/translation="MPRPFFHKLIPSTIQEKLRVPDKFVSKFDLSVAVALTVPD  
GHVWVGKADNKIWFQDQWQEFVDRYSIRIGVLLIFRYEGNSAPSVYIPNLSHSEI  
NYHSTGLMDSAHNHFKRALPEDLEDAEVIIPSSVYPSLPSTVPANKYASAI  
QTLTGVPKABEPTPTPKIPKRGKKKNADPEEINSSAPRDDDPENRSKFYESASAI  
KRTVTAEEERAINAAKTFEPTNPFVLRPSYLRGCIIMYLPSPGAPKYLISGSGF  
IKVQLAEQWQVPRCLYKAGRAKFSQWYEFTLENNLGRGDCVCFELLARTRDPVLKVT  
FRVNEYV"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,83e-139 Length: 1538  
Score: 1796.00 Matches: 341  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0  
US-10-088-187A-11 (1-341) x AF289052 (1-1538)  
Qy 1 MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20  
Db 312 ATGCCACGCCCTTCTTCATAAGTTGATTTTCTCATCCACTATCCAGAAACACGCTCG 371  
Qy 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40  
Db 372 AGGGTCCAGATAGTTTGTGAGTAAATTCAGGATGAGCTTTCGGTGTCTGTGACATC 431  
Qy 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60  
Db 432 ACAGTACCTGATGGTTCATGTTTGGCGTGTAGGACTAAGCAAGCTGCAACAAATTTGG 491  
Qy 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80  
Db 492 TTTCAAGATGGTGGCAAGAGTTTGTTCACCGTTACTCCATTCGATTCGTTATCTTTG 551  
Qy 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100  
Db 552 ATTTTGAATATGAAGAAACTCTGCCTTCAGCGTCTACATTTTCAATTTATCCACTCT 611  
Qy 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120  
Db 612 GAGATCAATTACCATTCCACCGGTCTCATGATTCGCTCCAGCTCAACCACTTCAACACGCC 671  
Qy 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr 140  
Db 672 CGTTTGTGAAGACCTTTGAAGATGAAGATCCGAGGTTCATCTTTCTCTCTCTGTAC 731

QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160  
DB 732 CCATCACCACCTTCTCTAGGCTACAGTACCAGCCAAAGAGGTATGCTAGTTCAGCCATC 791  
QY 161 GlnThrLeuPheThrGlyProValLysAlaGluGluProThrProThrProLysIlePro 180  
DB 792 CAAACCTTGTTCACCTGGACCACTTAAAGCTGAAGAGCCAAACGCCCAACCAATACCT 851  
QY 181 LysLysArgGlyArgLysLysLysAlaAspProGluGluIleAsnSerSerAlaPro 200  
DB 852 AAAAAGAGAGGAGGAGGAAGAAATAATGCTGATCTCGAGGAATAAATCTCATAGCTCCG 911  
QY 201 ArgAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220  
DB 912 CGAGATGATGATCCAGAGAACCGTTCAAGATTCTACGAGAGTCTTTCGCGAAGAGAGA 971  
QY 221 ThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThr 240  
DB 972 ACCGTGACTGCAGAAAGAGAGAGAGCCATCAATGCAGCCAAACGTTGCAACCAACA 1031  
QY 241 AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260  
DB 1032 AACCCCTTCTTCAGAGTGTCTCGACCATCTATCTATACAGAGGTGTCATCATGTAT 1091  
QY 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGln 280  
DB 1092 CTTCTCTCTGGTTGCTGAGAGTACCTTAAGTGGATCTCGGTTTCATCAAGTCCAG 1151  
QY 281 LeuAlaGluLysGlnTrpProValArgCysLysLysLysAlaGlyArgAlaLysPheSer 300  
DB 1152 CTTGGGAGAAACATGGCTGTTCGATGTCTCTACAAAGCGGAGAGCCAAATTCAGT 1211  
QY 301 GlnGlyTyrTyrGluPheThrLeuGluAsnLeuGlyGluGlyAspValCysValPhe 320  
DB 1212 CAAGGATGTGTACCAATTCTCTAGAGAACCACTTAGGAGAGGAGAGCTGTGTGTGT 1271  
QY 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyr 340  
DB 1272 GAGCTGCTCAGAACCCAGAGATTTCGTTTGAAGTGACAGCCCTTTCGAGTCAACCGATAC 1331  
QY 341 Val 341  
DB 1332 GTC 1334

RESULT 4  
AX101038 LOCUS 1495 bp DNA linear PAT 10-APR-2001  
DEFINITION Sequence 12 from Patent WO0121822.  
ACCESSION AX101038  
VERSION AX101038.1 GI:13619894

KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
1

REFERENCE Dean,C. and Levy,Y.Y.  
AUTHORS Methods and means for modification of plant flowering  
TITLE Characteristics  
JOURNAL Patent: WO 0121822-A 12 29-MAR-2001;  
Plant Bioscience Limited (GB)

FEATURES  
source  
1. .1495  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="vrnl-1 mutation"

ORIGIN  
Alignment Scores: 4.77e-138 Length: 1495  
Pred. No.: 1781.00 Matches: 340  
Score: 99.71% Conservative: 0  
Percent Similarity: 99.71% Mismatches: 1  
Best Local Similarity: 99.71%

Query Match: 99.16% Indels: 0  
DB: 6 Gaps: 0  
US-10-088-187a-11 (1-341) x AX101038 (1-1495)  
QY 1 MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20  
DB 269 ATGCCACGCCCTTCTTCCATAGTTGATTTCTCATCCACTATCCAAAGAAAACGCTG 328  
QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40  
DB 329 AGGTCCTCCAGATAAGTTGTGAGTAAATTCAGGATGAGCTTTCGGTGTCTGTTCACATC 388  
QY 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60  
DB 389 ACAGTACTGATGTCTGATTTGACGTGAGGTAAGGAAAGCTGACACAAATTTGG 448  
QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80  
DB 449 TTTCAAGATGGTTGGCAAGAGTTTGTTCAGCGTTTACTTCCATTCGATTTGGTATCTTTTG 508  
QY 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100  
DB 509 ATTTTATAGATATGAAGGAAACTCTGCCTTCAGCGTCTACATTTTCAATTTATCCACTCT 568  
QY 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120  
DB 569 GAGATCAATTACCATTCCACCGGTCTCATGGATTCCGCTCACCAACCACTTCAACCGGCC 628  
QY 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr 140  
DB 629 CGTTTGTGTAAGACCTTGAAGATGAAGATGCCGAGGTTCATCTTCTCTCTGTGTAC 688  
QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160  
DB 689 CCATCACCATCTCTCTAGTCTACAGTACAGCCAAAGAGGTATGCTAGTTCAGCCATC 748  
QY 161 GlnThrLeuPheThrGlyProValLysAlaGluGluProThrProThrProLysIlePro 180  
DB 749 CAAACCTTGTTCCTGGACCATTTAAAGCTGAAGAGCCACGCCCAACCCCAAAATACCT 808  
QY 181 LysLysArgGlyArgLysLysLysAlaAspProGluGluIleAsnSerSerAlaPro 200  
DB 809 AAAAAGAGAGGAGGAGAGAGAAATAATGCTGATCTCTAGGAAATAAATCTCATCAGCTCCG 868  
QY 201 ArgAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220  
DB 869 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGAGTGTCTTTCGCGAAGAGAGA 928  
QY 221 ThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThr 240  
DB 929 ACCGTGACTGCAGAAAGAGAGAGAGAGCCATCAATGCAGCCAAACGTTCCAAACCAACA 988  
QY 241 AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260  
DB 989 AACCCCTTCTTCAGAGTGTTCGCGACCATCTCTATATACAGAGGTGTCATCATGTAT 1048  
QY 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGln 280  
DB 1049 CTTCTCTCTGGTTGCTGAGAGTACCTTAAGTGGATCTCCGGTTTCATCAAGTCCAG 1108  
QY 281 LeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSer 300  
DB 1109 CTTGGGAGAAACAAATGGCTGTTCGATGTCTCTACAAAGCGGAGAGCCAAATTCAGT 1168  
QY 301 GlnGlyTyrTyrGluPheThrLeuGluAsnLeuGlyGluGlyAspValCysValPhe 320  
DB 1169 CAAGGATGGTACGAATTCACCTCTAGAGAACCACTTAGGAGAGGAGAGCTCTGTGTGT 1228  
QY 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyr 340  
DB 1229 GAGCTGCTCAGAACCCAGAGATTTCGTTTGAAGTGACAGCCCTTTCGAGTCAACGAGTAC 1288

```

QY      341 Val 341
Db      1289 GTC 1291

RESULT 5
AX101039
LOCUS   AX101039
DEFINITION Sequence 13 from Patent WO0121822.
ACCESSION AX101039
VERSION  AX101039.1 GI:13619895
KEYWORDS synthetic construct
SOURCE   synthetic construct
ORGANISM artificial sequences.
1
REFERENCE Dean,C. and Levy,Y.Y.
AUTHORS   Methods and means for modification of plant flowering
TITLE     Characteristics
JOURNAL   Patent: WO 0121822-A 13 29-MAR-2001;
          Plant Bioscience Limited (GB)
FEATURES Location/Qualifiers
          1. .1494
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="vrn1-2 mutation"

ORIGIN
Alignment Scores:
Pred. No.:      5.68e-137      Length:      1494
Score:          1768.00      Matches:    340
Percent Similarity: 99.71%      Conservative: 0
Best Local Similarity: 99.71%      Mismatches: 1
Query Match:     98.44%      Indels:    1
DB:              6          Gaps:      0

US-10-088-187A-11 (1-341) x AX101039 (1-1494)

QY      1 MetProArgProPheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20
Db      269 ATGCCACGGCCCTTCTTCATTAAGTGTGATTTCTCATCCACTATCCAGAAAACGTCG 328

QY      21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
Db      329 AGSGTCCACAGATAAGTTTGTGAGTAAATCAAGGATGAGCTTCGTTCTGTCACATC 388

QY      41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60
Db      389 ACAGTACCTCGTGGTGCATGTTTGGCGGTGAGGACTAAGGAAAGCTGACAAACAAATTTGG 448

QY      61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80
Db      449 TTTTCAGATGGTTGGCAGAGTTTGTGACCGTTACTCCATTCGCATTCGATTCCTTTTG 508

QY      81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100
Db      509 ATTTTATAGATATGAGGAAACTCTGCTTCAGCGTCTACATTTTCAATTTATCCCACTCT 568

QY      101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
Db      569 GAGATCAATTTACCATTTCCACCGGCTCTCATGGATTCGCTTCACACCACTTCAAAACGGCC 628

QY      121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr 140
Db      629 CGTTTGTGTTGAAGACCTTGAAGATGAAGATGCGGAGGTCACTTCCTTCCTTCGTGTGATC 688

QY      141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160
Db      689 CCATCACCACCTTCTCGAGTCTACAGTACCGACCAACAAAGGATGATGTAGTTCAGCCATC 748

QY      161 GlnThrLeuPheThrGlyProValLysAlaGluGluProThrProThrProLysIlePro 180
Db      749 CAAACCTTGTCTACTGGACCACTTAAAGCTGAAAGCCCAACGCCCAACCCCAAAAATACCT 808

```

```

QY      181 LysLysArgGlyYargLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaPro 200
Db      809 AAAAAGAGAGGGAGGAAGAGAA-AATGCTGATCTCTGAGGAATAAACTCATCAGCTCCG 867

QY      201 ArgAspAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220
Db      868 CGAGATGATGATCCAGAGAACGTTCAAAGTTCTACGAGAGTGTCTTCGCGAGAAAGAGA 927

QY      221 ThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThr 240
Db      928 ACCGTGACTGTCAGAAAGAGAGAGAGCCATCAATCGACGCCAAACAGTTCGAACCAACA 987

QY      241 AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260
Db      988 AACCCCTTCTTCAGAGTGGTCTCGGACCATCTCTATATACAGAGTTGCATCATGTAT 1047

QY      261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGln 280
Db      1048 CTTCTCTTCTGGGTTTCTGCTGAGAGTACCTAAGTGGGATCTCGGGTTCATCAAGTCCAG 1107

QY      281 LeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSer 300
Db      1108 CTTGCGGAGAAACAATGGCTCTTCGATGCTCTCAAAAGCCGGAGAGACCAATTCAGT 1167

QY      301 GlnGlyTrpTyrGluPheThrLeuGluAsnLeuGlyGluGlyValCysValPhe 320
Db      1168 CAAGGATGGTAGCAATTCACCTCTAGAACCACTTAGGAGAGAGAGACCTCTGTGTGTTT 1227

QY      321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyr 340
Db      1228 GAGCTGCTCAGAACCCAGAGATTTCGTTTGAAGTGTGACAGCCTTTCGAGTCAACGAGTAC 1287

QY      341 Val 341
Db      1288 GTC 1290

RESULT 6
AF289051      5000 bp      DNA      linear      PLN 05-NOV-2003
LOCUS   Arabidopsis thaliana reduced vernalization response 1 (VRN1) gene,
DEFINITION VRN1-Ler allele, complete cds.
ACCESSION   AF289051
VERSION     AF289051.1 GI:21734793
KEYWORDS   Arabidopsis thaliana (thale cress)
SOURCE     Arabidopsis thaliana
ORGANISM   Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1 (bases 1 to 5000)
AUTHORS     Levy,Y.Y., Menesse,S., Mylne,J.S., Gendall,A.R. and Dean,C.
TITLE       Multiple roles of Arabidopsis VRN1 in vernalization and flowering
JOURNAL     Science 297 (5579), 243-246 (2002)
MEDLINE     22111275
PUBMED      12114624
REFERENCE   2 (bases 1 to 5000)
AUTHORS     Levy,Y.Y., Gendall,A.R. and Dean,C.
TITLE       VRN1, a gene required for response to vernalization
JOURNAL     Unpublished
REFERENCE   3 (bases 1 to 5000)
AUTHORS     Levy,Y.Y. and Dean,C.
TITLE       Direct submission
JOURNAL     Submitted (23-JUL-2000) Molecular Genetics, John Innes Centre,
          Colney Lane, Norwich, Norfolk NR4 7UH, UK
FEATURES   Location/Qualifiers
            1. .5000
              /organism="Arabidopsis thaliana"
              /mol_type="genomic DNA"
              /db_xref="taxon:3702"
              /chromosome="3"
              /map="between mi339 and m560B2"

```

gene	ecotype="Landsberg erecta"
1879. .>4370	
/gene="VRN1"	
/allele="Ler"	
join(<1879. .1938,3120. .3567,3664. .3732,3832. .4031,	
4122. .>4370)	
/gene="VRN1"	
/allele="Ler"	
join(1879. .1938,3120. .3567,3664. .3732,3832. .4031,	
4122. .>4370)	
/gene="VRN1"	
/allele="Ler"	
/note="contains two VP1/ABI3-like B3 domains"	
/codon_start=1	
/product="reduced vernalization response 1"	
/protein_id="AAM76972.1"	
/db_xref="GI:21734794"	
/translation="MPRPFFKHLLIFSSSTIOEKRLRVPDKFVSKOELSVAVALTPVD	
GHVAVGKADNKIWFQDQWQEFVDYRISIRIGYLLIFRVEGNSAFVSIVIFNLHSEI	
NYHSGLDMSAHNHFKEARLPEDLEDEDAEVIIPSSVYPSLPSTVPANKGVASAI	
QTLFTGPVKAHEPTPTPKIPKRGKKKADPEINSSAPRDODDPENRSKFVESAR	
KRTVTAERERAINAKTFEPTFFRVLRPSLYIRGCIWYLPSPFAEKYLSGISGF	
IKVQLAEQWQVPRCLYKAGRAKFGQWYEFTEFLNLEGGDVCVFELLRDFVLKVTA	
FRVNEYV"	
ORIGIN	
Alignment Scores:	
Pred. No.:	1,31e-123
Score:	Length: 5000
Percent Similarity:	Matches: 321
Best Local Similarity:	Conservative: 0
Query Match:	Mismatches: 0
DB:	Indels: 95
	Gaps: 3
US-10-088-187A-11 (1-341) x AF289051 (1-5000)	
Qy	21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
Db	3120 AGGGTCCAGATAAGTTTGTGAGTAAATTCAGAGATGAGCTTTCGGTTGCTGTGTGCACTC 3179
Qy	41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60
Db	3180 ACAGTACCTGATGGTATGTTTGGCGTGTAGAGCTTAGGAAGCTGACACACAAATTTGG 3239
Qy	61 PheGlnAspGlyTrpGlnGluPheValAspArgTrpSerIleArgIleGlyTrpLeuLeu 80
Db	3240 TTTCAAGATGGTTGGCAAGATTTGTTGACCGTTACTCCATTCGCATTCGTTATCTTTTG 3299
Qy	81 IlePheArgTrpGluGlyAsnSerAlaPheSerValTrpIlePheAsnLeuSerHisSer 100
Db	3300 ATTTTATAGATAGAAGAACTCTGCGCTTCAGCGTCTACATTTTCAATTTATCCACATCT 3359
Qy	101 GluIleAsnTrpHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
Db	3360 GAGATCAATACCATTCACCGCTCATGGATTCGGCTCAACACCATTCAAACGCGCC 3419
Qy	121 ArgLeuPheGluAspLeuGluAspAlaGluValIlePheProSerSerValTrp 140
Db	3420 CGTTTGTTCGAACACCTTGAAGATGAGATGCGGAGTGCATCTTCTCTCTGTGTATC 3479
Qy	141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTrpAlaSerSerAlaIle 160
Db	3480 CCAATCCACCTTCTGTAGTCTACGATCCAGCCACAAAGGGTATGCTAGTTCAGCCATC 3539
Qy	161 GlnThrLeuPheThrGlyProValLys----- 169
Db	3540 CAAACCTTGTCTACGTGACCAAGTAAAGGTGATATTATAACCAACATGATTCCTTTATC 3599
Qy	169 ----- 169
Db	3600 TATCGCTGATACCGCTTATCATCTCTTTGAGGTGATGCTGTGATATTTTCCCTATCT 3659
Qy	170 ---AlaGluGluProThrProThrProLysIleProLysLeuArgGlyArgLysLysLys 188

## COMMENT

Tel:81-438-52-3935, Fax:81-438-52-3934)  
Address for correspondence: kaos@kazusa.or.jp  
For the latest information on annotation of this clone, please see  
http://www.kazusa.or.jp/kaos/cgi-bin/agd/graph.cgi?c=K13E13  
Genes with similarity to proteins in the databases are described in  
'product' or 'note' qualifiers. Genes that have no significant  
protein similarity are described as 'unknown protein'.  
The software programs used to predict genes include: Grail  
(Informatics Group, Oak Ridge National Laboratory,  
http://combio.ornl.gov/Grail-1.3/),  
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),  
NetGene2 (S.M. Hebegaard, et al., CBS Technical University of  
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and  
SplicePredictor (Volker Brendel, Stanford University,  
http://grenlini.zoool.iastate.edu/cgi-bin/sp.cgi).  
Genes encoding tRNAs are predicted by tRNAscan-SE  
(Sean Eddy, Washington University School of Medicine, St. Louis,  
http://genome.wustl.edu/eddy/tRNAscan-SE/).  
This sequence may not be the entire insert of this clone. It may be  
shorter because we remove overlaps between neighboring submissions.  
The 5' clone is MCB22 and the 3' clone is MHP21.

## FEATURES

## source

1. .79186  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3702"  
/chromosomes="3"  
/clone="K13E13"  
/clone\_lib="Miteui TAC"  
/ecotype="Columbia"  
392. .1558  
/note="unnamed protein product; gb|AAC69119.1  
gene id:K13E13.1  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01688.1"  
/db\_xref="GI:9280309"  
/translation="MKTIOEQPLNDLVEILCRVPATSLRLRSTCKAWNRLFKGDRI  
LASHPEKSAKQSLGSLNDVYRIFPISNHLGNSPISLEKSELIDPHSKNSAAPFEI  
SRVHCEGALLCSOLDSESRVVMPLTGETWRTGDPFROKGRSFDVGYVYQDKRS  
WIKSKYLCYRGYKFEIYDESDSWRLDIIAPRSGISYSELVSLSKNTYFAK  
GYTERPDTLGLKFDYTERKSPVLLPYQSRRLQASLSVVRDQKLVLLQLDQSS  
KTEIWNVDITETKAGSWTKVTLALDLSPHLQIGNDGSFPLGDKVVMFCEKLE  
NKVKDMVTVGDNVTEVGFVDMGCRVILNVPVSLVQIERAGGNRKRGH"  
complement (2742. .3977)  
/note="unnamed protein product; gb|AA01602.1  
gene id:K13E13.2  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01689.1"  
/db\_xref="GI:9280310"  
/translation="MRFVAPPPRSGNSPSPSGISSEILSRSSDPLPFPSPPLIA  
MVVLAAPLVTVYGRSLRFLPLFRFRWRCKRRLRLHLSASSASTSSDLRS  
FPFPPDFSHYSYSPYGLDDSVIKTLPFLYAAACGKPAVGKTSAAACRDCAVCL  
LFEFGDYVTLPLCFHAFLECIDELRSHENCPLCRTLIIAGSAGVLTWSPFVPLM  
APRIPSDIDENNAIITRGETTPSRNWTIAADTNDQEIIRASVEQSPAISRFPR  
ELKRSVPCECESSESERVMEPATVSPWRVRSRTWNRQSPGNLISKSVFSERY  
RSTKSPFFRRSSAGVFPISERIIPATGSSSRRTKSTSPMPFFRTAPHSRSLRCGDP  
EALLSPERRRRDTCRAEN"  
complement (Join (5713. .5792,5883. .6006,6112. .6349,  
6579. .6697,6873. .6931,7161. .7326,7623. .7682))  
/note="unnamed protein product; gene\_id:K13E13.3  
unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01690.1"  
/db\_xref="GI:9280311"  
/translation="MEFVVEEGLVNEDCSTLVLPAISIGNVQLAVDLLVSSSTAER  
VGYLDDPPLLPCVNDAYGPLPCGIALPLEVYESSSTATTLAQORSVPAGMKIKFA  
ENIAFASSGKKHIIIVSSLDLQRLHNDWSRGQVYLLSNAESDGDHCHERIGFG  
RLHEDYSEGRCKWYLLSSVPEKNSVBEALPSEDELEIDYYPSPFLAALFAFKARGL

## CDS

/evidence=not experimental  
join (21250. .21381,21635. .21952,22001. .22061,22130. .22200)  
/note="unnamed protein product; gb|AAC72857.1  
gene id:K13E13.7  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01692.1"  
/db\_xref="GI:9280313"  
/translation="MVTQNTKARVTSVSHRSQQDPSPVKPFKFLVLPSTMKDKMM  
KIPRFVKLQGSLSSEVLTETPAGFKRSIKLGRIGERIWFHEGHSFAEASIEGH  
FLLEPKNSFRVLIIPNVSACTKYPLDAVHIIDSDDDIIDITGKEPDETHSKKRS  
RDIELDKLHLDVWNPVHLKESEEDKRVFRG"  
complement (22331. .23749)  
/note="unnamed protein product; gb|AAF15933.1  
gene id:K13E13.8  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01693.1"  
/db\_xref="GI:9280314"  
/translation="MSSVFPGRFLSLQONSKTLLQAKIHAQLVINGCHDNSLFGK  
LIHYCSKPTSESSKLAHLVFPFRGHPDKFLFNTLLCKSPEDSIRIFANYASKSS  
LLYLNERTFVLGACASASSALRVGRIVHGVMVKGLFVSELSIGTLLHLYAKN  
GDLRYARKVDEMPERTSVTNWMLGGYCSHDKGNHARAKMVLPRFSCGSGVRP  
TDTTWCVLISAISQTLGELISLVHGYTEKLGFTPEVDVFIGTALVDMYKCGCLNNA  
FSVLEMLKRVNFTNATGLANGRNETPNLNRMAESGINKNETITFTSLLSAYR  
HIGVLEGIELFKMTRFGVTPVIEHYGCTVLLGKAGRIQEAQFTILAMPIDKDAI  
LLRLSNACYSYGETWGEIEGKLLLETERDEKLSGSECEYVALSVLAHKGKWE  
VEKLKMKERRIKTRPGYSFV"  
25146. .26393  
/note="unnamed protein product; gb|AAC69119.1  
gene id:K13E13.9  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01694.1"  
/db\_xref="GI:9280315"  
/translation="MTIPDLNCNDLVEILCRVPARNLKRSLRSTKRWNLFPKDDRRFA  
REHMKAPKEYLPLMLSEYRIPCIVNSLQDVPVSLKRELSLPDPYQFDTGRV  
FHCGLGLVWGNKPNRYGSKI VVWNPVTGTGTWTEAGYRWKEVRFVGLVGYCOQDE  
NNSCSKKYKTLCPYPNGQDTIEIYELNYSRWRTIPDGLTIPGWTLYISQTYSMNG  
NLYIFASKSKPHGLVSLRDFSTKSSCLVTLIPQPRIYEIISLAVRGENJULL  
LOLDFSEKTEITWNKIDTTTKGAASVTKVLAPDLSPLQLFSEEVNLFLEDDKKY  
AVCCERMLPEQEHRYQCRREYKITDKTIYILGSDNKNVDEVGSGEGEATDLSLEGSQVI

## CDS

/evidence=not experimental  
/protein\_id="BAB01694.1"  
/db\_xref="GI:9280315"  
/translation="MTIPDLNCNDLVEILCRVPARNLKRSLRSTKRWNLFPKDDRRFA  
REHMKAPKEYLPLMLSEYRIPCIVNSLQDVPVSLKRELSLPDPYQFDTGRV  
FHCGLGLVWGNKPNRYGSKI VVWNPVTGTGTWTEAGYRWKEVRFVGLVGYCOQDE  
NNSCSKKYKTLCPYPNGQDTIEIYELNYSRWRTIPDGLTIPGWTLYISQTYSMNG  
NLYIFASKSKPHGLVSLRDFSTKSSCLVTLIPQPRIYEIISLAVRGENJULL  
LOLDFSEKTEITWNKIDTTTKGAASVTKVLAPDLSPLQLFSEEVNLFLEDDKKY  
AVCCERMLPEQEHRYQCRREYKITDKTIYILGSDNKNVDEVGSGEGEATDLSLEGSQVI

## CDS

/evidence=not experimental  
/protein\_id="BAB01694.1"  
/db\_xref="GI:9280315"  
/translation="MTIPDLNCNDLVEILCRVPARNLKRSLRSTKRWNLFPKDDRRFA  
REHMKAPKEYLPLMLSEYRIPCIVNSLQDVPVSLKRELSLPDPYQFDTGRV  
FHCGLGLVWGNKPNRYGSKI VVWNPVTGTGTWTEAGYRWKEVRFVGLVGYCOQDE  
NNSCSKKYKTLCPYPNGQDTIEIYELNYSRWRTIPDGLTIPGWTLYISQTYSMNG  
NLYIFASKSKPHGLVSLRDFSTKSSCLVTLIPQPRIYEIISLAVRGENJULL  
LOLDFSEKTEITWNKIDTTTKGAASVTKVLAPDLSPLQLFSEEVNLFLEDDKKY  
AVCCERMLPEQEHRYQCRREYKITDKTIYILGSDNKNVDEVGSGEGEATDLSLEGSQVI

CDS  
LNYPASLVQIEQAGGKTKRGDD"  
complement (join(27709..27957,28048..28247,28347..28409,  
28512..28692))  
/note="unnamed protein product; gb|AAD43153.1  
gene\_id:K13E13.10  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01695.1"  
/db\_xref="GI:9280316"  
/translation="MDSHHNFKEARLPEDLEDAEVIKPSVVPSPLESTVPANK  
GYASAIQTLFGPKVETPTPKPKGKRKKKNADPEEINSSAPRDDDDENRSEKYE  
SASAKRTVTAERERAINAAKTEPTFPFRVIRFSILYRGCMILYPSGFAEKYLS  
GISGFIKVLAKQWPKVRLYKAGRAKFSQGWYFTLENNIGDGVCFELLRTDFV  
LKVTAFRVNEV"  
complement (join(32300..32554,32669..32996,33144..33619))  
/note="gene\_id:K13E13.11"  
/codon\_start=1  
/evidence=not experimental  
/product="oxylase-like protein"  
/protein\_id="BAB01696.1"  
/db\_xref="GI:9280317"  
/translation="MGELDEAFIOAEPHRPNTHLTNSGDFIFSDIPIIDLSLEDTH  
HKTALEIAEACKWGFQVNHGLPSALHRVETAAEFNLTTEKKRVKRDVE  
NMGVHDEHTKNVDWEIEFDPLQDSTIVPASPEPEDTELKLTNQWPNPSPHRE  
VCOEYAREVEKLAFLLELVLISLGLPGRUTFGNEQTSFLRFNHYPCPNPLALG  
VGRHKDGGALTYLAQDSVGLQVRSRSDQWIPVKPISDALIINMGNCIQWTDNEYW  
SABHRVNVNTSKERFSIPFFPPSHSEANIEPLEELISEENPPCYKNWGWFFVSRNR  
SDPKLEVENIQIDHFKA"  
complement (join(34938..35198,35304..35631,36034..36494))

Alignment Scores:  
Pred. No.: 3,98e-122 Length: 79186  
Score: 1614.50 Matches: 321  
Percent Similarity: 77.16% Conservative: 0  
Best Local Similarity: 77.16% Mismatches: 0  
Query Match: 89.89% Indels: 95  
DB: 8 Gaps: 3

US-10-088-187A-11 (1-341) x AF000735 (1-79186)

QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40  
DB 28959 AGGTCCTCCAGATAAGTTGTGAGTAATTCAGGATGAGCTTCGGTTCGTGCTGCTC 28900

QY 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60  
DB 28899 ACAGTACCTGATGGTCATGTTGGCGGTGTAGGACTAAGGAAAGCTGACACAAATTTGG 28840

QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTrpSerIleArgIleGlyTrpLeuLeu 80  
DB 28839 TTTCAAGATGGTTGGCAGAGTGTGTGACCGTTACTCCATTCGCTGCTTATCTTTG 28780

QY 81 IlePheArgTrpGluGlyAsnSerAlaPheSerValTrpIlePheAsnLeuSerHisSer 100  
DB 28779 ATTTTAGATATGAAGAACTCTGCTTCAGGCTTACATTTTCAATTTATCCCACTCT 28720

QY 101 GluIleAsnTrpHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120  
DB 28719 GAGATCAATTAACCATTCACCGGCTCATGGATTCGCTCAACCACTTCAACCGCC 28660

QY 121 ArgLeuPheGluAspLeuGluAspAlaGluValIlePheProSerSerValTrp 140  
DB 28659 CGTTTGTGGAGACCTTGAAGATGAAGTCCGAGGTGATCTTCTCTCTCTGTGTAC 28600

QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTrpAlaSerSerAlaIle 160  
DB 28599 CCATCACCACCTCTCTGCTGCTACAGTACCAGCAACAAAGGTATGCTAGTTCCAGCCAT 28540

QY 161 GlnThrLeuPheThrGlyProValLys----- 169  
DB 28539 CAAACCTTGTTCACTGGACCAAGTTAAAGGTGATATTTATAACCAACTGATTCCTCTTATC 28480

QY 169 ----- 169  
DB 28479 TATCGCTGATTAGCGGCTTATCATCTTTTTCAGGTTGATGCTTATATTTTCCTTATCT 28420

QY 170 ---AlaGluGluProThrProThrProLysLysArgGlyArgLysLysLys 188  
DB 28419 CCAGCTGAAGAGCAACGCAACCCCAAAATACCTTAAAGAGAGGAGGAGGAGAA 28360

QY 189 AsnAlaAspPro----- 192  
DB 28359 AATGCTGATCTCGTAAAGCACTTTTCTTGAATGCTTCAGACTCGTTTTCAGAGA 28300

QY 193 -----GluGluLeu 195  
DB 28299 TTCACAGATCTCTCATGATACATATATCTTTTGAATGCTTCCTTACAGAGGAAATA 28240

QY 196 AsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTrpGluSerAla 215  
DB 28239 AACTCATCAGTCCGCGAGATGATGATCCAGAGAACCGTTCAAGTTCCTACGAGGTCT 28180

QY 216 SerAlaArgLysArgThrValThrAlaGluGluArgAlaAlaIleAsnAlaLys 235  
DB 28179 TCTGCGAGAAAGAAACCGTGACTGCGAGAAAGAGAGAGGCCATCATCGAGCAAA 28120

QY 236 ThrPheGluProThrAsnProphePheArgValValLeuArgProSerTrpLeuTrpArg 255  
DB 28119 ACGTTCGAACCAACCAACCTTTCTTTCAGAGTGGTTCCTGCGACCATCTCTATACAG 28060

QY 256 GlyCysIleMet----- 259  
DB 28059 GGTTGCATCATGTGTAATAAAAAACATCTTAGAGAGACTTAATCTTATCGGTCTCTCAC 28000

QY 260 -----TyrLeuProSerGlyPhe 265  
DB 27999 TGATCTCTAAAGAGCCCTTCTGTTCTGTTCTCTCAACAGATATCTCTCTCTGCGTTT 27940

QY 266 AlaGluLysTrpLeuSerGlyLysPheIleLysValGlnLeuAlaGluLysGln 285  
DB 27939 GCTGAGAAGTACCTTAAGTGGGATCTCGGGTTTCATCAAGTCCAGCTTGGCGGAAACAA 27880

QY 286 TrpProValArgCysLeuTrpLysAlaGlyArgAlaLysPheSerGlnGlyTrpTrpGlu 305  
DB 27879 TGGCTGTTCATGCTCTTACAAAGCGGAGAGCCAAATTCAGTCAAGGATGGTACGAA 27820

QY 306 PheThrLeuGluAsnLeuGlyGlyAspValCysValPheGluLeuLeuArgThr 325  
DB 27819 TTCACTCTAGAGAACAACTTAGGAGAGAGAGAGCTCTGTGTGTGTGAGCTGCTCAGAAC 27760

QY 326 ArgAspPheValLeuLysValThrAlaPheArgValAsnGluTrpVal 341  
DB 27759 AGAGATTTGCTTTTGAAGTGACAGCCCTTCGAGTCAACGAGTACGTC 27712

RESULT 8  
AX101027 5000 bp DNA linear PAT 10-APR-2001  
LOCUS Sequence 1 from Patent WO0121822.  
DEFINITION AX101027  
ACCESSION AX101027  
VERSION AX101027.1 GI:13619883  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
1  
REFERENCE  
AUTHORS Dean,C. and Levy,Y.Y.  
TITLE Methods and means for modification of plant flowering  
JOURNAL characteristics  
PATENT Patent: WO 0121822-A 1 29-MAR-2001;  
LOCATION/Qualifiers Plant Bioscience Limited (GB)  
1..5000  
/organism="Arabidopsis thaliana"

/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

## ORIGIN

Alignment Scores:  
Pred. No.: 1.56e-122 Length: 5000  
Score: 1601.50 Matches: 320  
Percent Similarity: 76.92% Conservative: 0  
Best Local Similarity: 76.92% Mismatches: 1  
Query Match: 89.17% Indels: 95  
DB: 6 Gaps: 3

US-10-088-187A-11 (1-341) x AX101027 (1-5000)

QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40  
DB 3120 AGGGTCCCAGATAAGTTTGTGAGTAATTCAGAGATGAGCTTCGGTTCGTGTCACATC 3179

QY 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60  
DB 3180 ACAGTACCTGATGGTTCATGTTTGRCGTGTAGGACTAAGGAAAGCTGACAAACAAATTTGG 3239

QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrIleuLeu 80  
DB 3240 TTTCAGATGGTTGGCAAGAGTTTGTGACCGTTACTCCATTGCGCATTCGTTATCTTTTG 3299

QY 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100  
DB 3300 ATTTTATGATATGAGGAACCTGCTTCAGGCTCTACATTTTCATTTATCCACATCT 3359

QY 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120  
DB 3360 GAGATCAATTACCATTCACCGCTCTCATGGATTCGCTCACAAACCATTCAAACGCGCC 3419

QY 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr 140  
DB 3420 CGTTTTTTTGAAGACCTTGAAGATGAAGATGCGAGGTTCATCTTCTCTTCTGTTGATC 3479

QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160  
DB 3480 CCATCACCACTTCCTGAGTCTACAGTACCAGCCACAAAGGATGCTAGTTTCAGCCATC 3539

QY 161 GlnThrLeuPheThrGlyProValLys 169  
DB 3540 CAAACCTTGTCTACTGGACCACTTAAAGGTGATATTTATACCAACTGATTCCTTTATTC 3599

QY 169 ----- 169

DB 3600 TATCGCTGATTACGGCTCTTATCATCTTTTGGAGTTGATGCTTGATATTTTCTTATCT 3659

QY 170 ---AlaGluGluProThrProThrProLysIleProLysLysArgGlyArgLysLysLys 188  
DB 3660 CCAGCTGAGAGCAACGCGCAACCCCAAAATACCTTAAAGAGAGGAGGAGGAAGAA 3719

QY 189 AsnAlaAspPro----- 192  
DB 3720 AATGCTGATCCTGGTAAGCACTTTCTCTTTGMAATGCTTCAGACTCGTTTTTCAGAGGA 3779

QY 193 -----GluIle 195  
DB 3780 TTTCACAGATTCTTCTCATGATACATATATCTTTTGTATATTTGCTTTCAGAGGAATA 3839

QY 196 AsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrGluSerAla 215  
DB 3840 AACTCATAGCTCCGCGAGATGATGATCCAGAGAACCGTTCAAGTTTCTACAGAGTGTCT 3899

QY 216 SerAlaArgLysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAlaLys 235  
DB 3900 TCTGCAGAGAAGAAACCGTACTGTCAGAGAAGAAGAGAGAGCCATCAATCAGCCAAA 3959

QY 236 ThrPheGluProThrAsnProPhePheArgValValLeuArgProSerTyrIleuTyrArg 255  
DB 3960 ACGTTTCGAACCAACCAACCTTTCTTTCAGAGTGGTTCTCGACCATCTCTATCTATACAGA 4019

QY 256 GlyCysIleMet----- 259  
DB 4020 GGTTCATCATGGTAATAAAAAACATCTTAGGAAGACTTAATCTTATCGGTGCTCTTCAC 4079

QY 260 -----TyrIleuProSerGlyPhe 265  
DB 4080 TGATCTCTAAAGAAGACCTTCTGTTTCTCTCTCAACAGTATCTCTCTCTCGGTTT 4139

QY 266 AlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGlnLeuAlaGluLysGln 285  
DB 4140 GCTGAGAAGTACTAAGTGGGATCTCCGGGTTTCATCAAGTCCAGCTTCGGAGAAACAA 4199

QY 286 TrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSerGlnGlyTrpTyrGlu 305  
DB 4200 TGGCCTGTTGATGTCTCTACAAAGCCGGAGAGCAAAATTTCAGTCAAGGATGGTACGAA 4259

QY 306 PheThrLeuGluAsnAsnLeuGlyGluGlyAspValCysValPheGluLeuLeuArgThr 325  
DB 4260 TTCACTCTAGAGAACAACCTTAGGAGAAGGAGACGCTCTGTGTTTGTAGCTGCTCAGAAC 4319

QY 326 ArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyrVal 341  
DB 4320 AGAGATTTCGTTTTGAAAGTGACGCTTTCGAGTCAACGAGTACGTC 4367

RESULT 9  
LOCUS AY517929 990 bp mRNA linear PLN 18-JAN-2004  
DEFINITION Brassica rapa cultivar Kwonsim reduced vernalization response 1  
ACCESSION AY517929  
VERSION AY517929.1 GI:40806798  
KEYWORDS  
SOURCE Brassica rapa (Brassica campestris)  
ORGANISM Brassica rapa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 990)  
AUTHORS Kwon,S.-J., Park,B.-S., Kim,S.-Y., Choi,H.-S., Lee,M.-C., Kim,J.-S., Lee,S.-I., Lim,K.-B., Kim,J.-A., Lee,M.-R., Jin,Y.-M., Kim,D. and Kim,H.-I.  
TITLE Brassica rapa Kwonsim VRN1 mRNA  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 990)  
AUTHORS Kwon,S.-J., Park,B.-S., Kim,S.-Y., Choi,H.-S., Lee,M.-C., Kim,J.-S., Lee,S.-I., Lim,K.-B., Kim,J.-A., Lee,M.-R., Jin,Y.-M., Kim,D. and Kim,H.-I.  
TITLE Direct Submission  
JOURNAL Submitted (05-DEC-2003) Genomics Division, National Institute of Agricultural Biotechnology, Seodun-dong 225, Suwon 441-707, Republic of Korea  
FEATURES  
source location/Qualifiers  
1..990  
/organism="Brassica rapa"  
/mol\_type="mRNA"  
/cultivar="Kwonsim"  
/db\_xref="taxon:3711"  
1..990  
/note="VRN1"  
/codon\_start=1  
/product="reduced vernalization response 1"  
/protein\_id="AAR92199.1"  
/db\_xref="GI:40806799"  
/translation="MPREFFPKLIFSSTIQSKRLVDPKFSRPFKDELSVALVTPD GHVWRVGLRKADNNKWFQDQWQEFVDYRSIRIGYLLIFRYGNSAFSVCILNLPQS EINTHSTGLMDSASHNNHFKRPLFEDLEDAETLHTTASAIQSFPTGPKVKEPATP TOTSKVPKKRGKKKNADHPESVNSAPRDDPSRSKFTVSARKKRTVNAERERA VNAAKTREPPTNPFRRVLRLPSLYRGCIIMVLPSPGPKYLSGISGFIKVLGEKQWPV RCLYKAGRAKFSQGWYETVENNLEGBDVCPELLTRDFVLKVTAIRVNEVY"

CDS  
1..990  
/note="VRN1"  
/codon\_start=1  
/product="reduced vernalization response 1"  
/protein\_id="AAR92199.1"  
/db\_xref="GI:40806799"

ORIGIN  
Alignment Scores:

Pred. No.:	4,51e-116	Length:	990
Score:	1513.00	Matches:	296
Percent Similarity:	88.73%	Conservative:	13
Best Local Similarity:	85.06%	Mismatches:	13
Query Match:	84.24%	Indels:	26
DB:	8	Gaps:	5
US-10-088-187A-11 (1-341) x AY517929 (1-990)			
QY	1	MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu	20
Db	1	ATGCCACGCCCTTCTTCACAAAGTTGATTTCTCATCCACTATCCAAAGAAAGCGTCTA	60
QY	21	ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu	40
Db	61	AGAGTTCAGATTAAGTTGTGATGATTCAGAGCAGACTATCGTTGCGATGTCATC	120
QY	41	ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAsp-----AsnLys	58
Db	121	ACAGTACCGGATGGTTCATGTTGGCGTGTAGGACTAAGGAAAGCTGACAAACAACAAG	180
QY	59	IleTrpPheGluAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyr	78
Db	181	ATTGTGTTTCAAGATGGTTGGCAAGAGTTGTGTGACCGTTACTCATCGGATGGTTAC	240
QY	79	LeuLeuIlePheArgTyrGluGlyAenSerAlaPheSerValTyrIlePheAsnLeuSer	98
Db	241	CTTTTGATCTTGTAGTACGAGGCACTCTGCTTCAGCGTGTGATTTACACTTACCA	300
QY	99	HisSerGluIleAsnTyrHisSerThrGlyLeuMetAspSerAla-----HisAsnHis	116
Db	301	CAGTCCGAGATTAACTACCATTCACCGGTCTCATGGACTCTGCATCACACAACACAC	360
QY	117	PheLysArgAlaArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePhePro	136
Db	361	TTCAAAACGTCGCCGTTGTTTGAAGACCTTGAAGATGAAGATGCTGAGACTCTGCACA	420
QY	137	SerSerValTyrProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAla	156
Db	421	ACC-----	423
QY	157	SerSerAlaIleGlnThrLeuPheThrGlyProValLysAlaGluGlu-----ProThr	174
Db	424	GCTTACGCCATCCAGAGCTTCTTCACTGACCTGTTTAAACCTGGAAGAGGCAACACA	483
QY	175	ProThrProLysIleProLysLysArgGlyArgLysLysAlaAsp---ProGlu	193
Db	484	CAAACTCAAAAGTTCTTAAAGAGAGGAGGAGAGAAAGAAATGCTGATCATCTGAG	543
QY	194	GluIleAsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrGlu	213
Db	544	GAAGTAAATCATCTGCACCGAGAGATGATGACCCAGAGCGCGTTCAAAGTTCTACGAG	603
QY	214	SerAlaSerAlaArgLysArgThrValThrAlaGluGluArgLysAlaIleAsnAla	233
Db	604	AGTGCTTCTCGGAGAAAGAGAACAGTGAATGCTGGAAGAGAGAGAGGGCTGTTAATGA	663
QY	234	AlaLysThrPheGluProThrAenProPheArgValValLeuArgProSerTyrLeu	253
Db	664	GCCAAACGTTTGAGCCAAACAACCTTCTTCAGAGTTGTTCTGAGACCGTCTTATCTA	723
QY	254	TyrArgGlyCysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyIle	273
Db	724	TACAGAGTTGTCATCATGATTTGCTTCTGTTTGTGAGAGTACTTAAGTGGGATC	783
QY	274	SerGlyPheIleLysValGlnLeuAlaGluLysGlnTrpProValArgCysLeuTyrLys	293
Db	784	TCGGGATTCATCAAGGTCAGGAGGAGAAACAATGGCCCTGTGAGATGCCATATACAA	843
QY	294	AlaGlyArgAlaLysPheSerGlnGlyTyrTyrGluPheThrLeuGluAsnLeuGly	313
Db	844	GCAGGAGAGCTAAGTTCAAGTTGATGATGATTTCACTGTGGAGATTAACCTAGGA	903
QY	314	GluGlyAspValCysValPheGluLeuLeuArgThrArgAspPheValLeuLysValThr	333
Db	904	GAAGTGATGTCGCGTGTGTTGAGCTGCTGAGAACAGAGATTCGTTTGAAGTGACG	963
QY	334	AlaPheArgValAenGluTyrVal	341
Db	964	GCCTATCGAGTCAACGAGTACGTC	987
RESULT 10			
LOCUS	AY356368	990 bp	mRNA linear
DEFINITION	Brassica rapa cultivar Samjin reduced vernalization response 1	PLN 25-AUG-2003	
ACCESSION	AY356368		
VERSION	AY356368.1	GI:33943516	
KEYWORDS	Brassica rapa		
SOURCE	Brassica rapa		
ORGANISM	Brassica rapa		
REFERENCE	1 (bases 1 to 990)		
AUTHORS	Kwon,S.-J., Park,B.-S., Kim,S.-Y., Choi,H.-S., Lee,M.-C., Kim,J.-S., Lee,S.-I., Lim,K.-B., Kim,J.-A., Hong,K.-Y., Lee,M.-R., Jin,Y.-M., Kim,D. and Kim,H.-I.		
TITLE	Submitted (23-JUL-2003) Genomics Division, National Institute of		
JOURNAL	Agricultural Biotechnology, Seodun-dong 225, Suwon 441-707, Republic of Korea		
FEATURES	Location/Qualifiers		
source	1..990		
	/organism="Brassica rapa"		
	/mol_type="mRNA"		
	/cultivar="Samjin"		
	/db_xref="taxon:3711"		
CDS	1..990		
	/note="VRN1"		
	/codon_start=1		
	/product="reduced vernalization response 1"		
	/protein_id="AA055453.1"		
	/db_xref="GI:33943517"		
	/translation="MPPPFKLIFFSTIOEKRLRVDPKFSKFDLSVAVALTPD GHVWRGLRKADNNKIKWFQDQHQEQFVDRYSIRIGYLLIFRYSGNSAFSVCIVNLPOS EINYHSTGLMDSHNNHFKRPLIFDEDEDAETLLITASAIQSFTGPKVEEATP TOTSKVKKRGRKKKNADHPENVSSAPRDDPESKSKFYESARSGTVMABERERA VNAKTEPTNPFRVVRPSLYGCMILPSCFAEKYLSGTSGLTKVQGEKQWPC RCLYKAGRAKFSQGWYEFTEVNNLGEDVCVFELLRTDRDFVLKVTAYRVNEYV"		
Alignment Scores:			
Pred. No.:	9.67e-116	Length:	990
Score:	1509.00	Matches:	296
Percent Similarity:	88.51%	Conservative:	12
Best Local Similarity:	85.06%	Mismatches:	14
Query Match:	84.02%	Indels:	26
DB:	8	Gaps:	5
US-10-088-187A-11 (1-341) x AY356368 (1-990)			
QY	1	MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu	20
Db	1	ATGCCACGCCCTTCTTCACAAAGTTGATTTCTCATCCACTATCCAAAGAAAGCGTCTG	60
QY	21	ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu	40
Db	61	AGAGTTCAGATTAAGTTGTGATGATTCAGAGCAGACTATCGTTGCGATGTCATC	120
QY	41	ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAsp-----AsnLys	58
Db	121	ACAGTACCGGATGGTTCATGTTGGCGTGTAGGACTAAGGAAAGCTGACAAACAACAAG	180
QY	59	IleTrpPheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyr	78



Db 170 AGCTCTTATGCAATGGATTCGGACAGATCAGTTCACAAACAGTCTGATTTGTA 229  
 QY 124 uAspLeuGluAspGluAlaGluValIlePheProSerValTyrProSerProLe 144  
 Db 230 AGATCTGAACTCAAGATGCTAAGTCAATATCATTCG-----AT 271  
 QY 144 uProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAla---lleGlnThrLe 163  
 Db 272 CCTCGAATCTACTGAACCACTGAATAAAGGTTATGGCGTTCTACAGACATCCAAAGCTT 331  
 QY 163 upheThrGlyProValLysAlaGluGluProThrProThrProLysLleProLysLysAr 183  
 Db 332 TTTTCAAA---GAATCTAAAGCTGAAGAA-----ACGCCAAGGCTTCTAGAGAG 379  
 QY 183 GGLyArgLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAspAs 203  
 Db 380 AGGAGGAGAGAGAGAGATCTTAATCCCGAGGAGTAACCTTCAACTCCCGGTGGAGA 439  
 QY 203 pAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrValTh 223  
 Db 440 TGACTCAGAGAACCGCTCAAGATCTTACGAGAGTCTTCTGCTGAAAGAGAACTGTAAAC 499  
 QY 223 rAlaGluGluArgGluArgAlaIleAsnAlaLysThrPheGluProThrAsnProPh 243  
 Db 500 TGCAGAGGAGAGAGAGAGAGCGCTCAATCGACCAACATTCGACCAACATCTTA 559  
 QY 243 ePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysLleMetTyrLeuProSe 263  
 Db 560 CTTTAGAGTTGTTCTCGCAACCATCATATATACAGAGGTTGCATCATGTACTTCCCATC 619  
 QY 263 rGlyPheAlaGluLysTyrLeuSerGlyLysPheLysValGlnLeuAlaGl 283  
 Db 620 TGGGTTTGTCTGAAATACCTAAGTGGATATCTGTTTCATCAGCTCCAGCTCGTGA 679  
 QY 283 uLysGlnTyrProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSerGlnGlyTr 303  
 Db 680 GAAACATGCGCAGTGGTGCTCTACAAGAGGAGAGAGTAAAGTTAGCCAAAGGATG 739  
 QY 303 pTyrGluPheThrLeuGluAsnLeuGlyGluGlyAspValCysValPheGluLeuLe 323  
 Db 740 GTATGATTCACACTCGAGAACAAATATAGCGAGAGAGATGTATGTGTTTGAGCTACT 799  
 QY 323 uArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyrVal 341  
 Db 800 CAGAACTCGGGATTCGTTCTCGAAGTCCAGCTCCGCTTCTGTCATGATGATGTG 854  
 RESULT 12  
 AC007504 125021 bp DNA linear PLN 30-OCT-2002  
 LOCUS Arabidopsis thaliana chromosome I BAC F13F21 genomic sequence,  
 DEFINITION complete sequence.  
 VERSION AC007504  
 ACCESSION AC007504.3 GI:5430745  
 KEYWORDS HTG.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 125021)  
 Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.  
 REFERENCES  
 AUTHORS  
 1 (bases 1 to 125021)  
 Federespiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,  
 Altafi, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,  
 Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,  
 Lueros, S., Schwartz, J., Shinn, P., Toriumi, M., Vyotskaia, V.,  
 Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.  
 JOURNAL  
 REFERENCES  
 AUTHORS  
 2 (bases 1 to 125021)  
 Federespiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,  
 Altafi, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,  
 Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,  
 Lueros, S., Schwartz, J., Shinn, P., Toriumi, M., Vyotskaia, V.,  
 Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 FEATURES  
 source

Direct Submission  
 Submitted (06-MAY-1999) DNA Sequencing and Technology Center,  
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
 USA  
 3 (bases 1 to 125021)  
 Federespiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,  
 Altafi, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,  
 Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,  
 Lueros, S., Schwartz, J., Shinn, P., Toriumi, M., Vyotskaia, V.,  
 Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.  
 Direct Submission  
 Submitted (08-JUN-1999) DNA Sequencing and Technology Center,  
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
 USA  
 4 (bases 1 to 125021)  
 Federespiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,  
 Altafi, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,  
 Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,  
 Lueros, S., Schwartz, J., Shinn, P., Toriumi, M., Vyotskaia, V.,  
 Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.  
 Direct Submission  
 Submitted (09-JUL-1999) DNA Sequencing and Technology Center,  
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
 USA  
 On Jul 9, 1999 this sequence version replaced gi:5019265.  
 e-mail for correspondence: arab@sequence.stanford.edu  
 Genes with similarity to proteins in the databases are described  
 as 'putative', '-like' or 'similar to'. Genes that have EST  
 similarity but no significant protein similarity are described as  
 'unknown proteins'. Genes that are annotated based only on gene  
 prediction software are described as 'hypothetical proteins'.  
 The software programs used to predict genes include: Grail  
 (Informatics Group, Oak Ridge National Laboratory,  
 http://compbio.ornl.gov/section/index.html), GENSCAN (Chris Burge,  
 http://gnomic.stanford.edu/~chris/GENSCANW.html), Fexa (V. Solovyev  
 & A. Salamov, Sanger Centre, http://genomic.sanger.ac.uk/),  
 NetPlantGene (S.M. Hebsgaard, et al., CBS, Technical University of  
 Denmark, http://www.cbs.dtu.dk/NetPlantGene.html) and  
 eMotif (Nevill-Manning, C.G. Wu, T.D. & Brutlag, D.L.,  
 http://motif.stanford.edu/projects.html).

1. 125021  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3702"  
 /chromosome="I"  
 /map="mi208"  
 /clone="F13F21"  
 complement(4079..5020)  
 /gene="F13F21.1"  
 complement(join(4079..4465,4631..5020))  
 /gene="F13F21.1"  
 /note="Hypothetical Protein"  
 /codon\_start=1  
 /protein\_id="AAD43146.1"  
 /db\_xref="GI:5430746"  
 /translation="MTNLISGSSPTTANLTFPMQWKIECWLRHAEPSVDETICOMVEI  
 MTLFERYWEYSDILAIAVDFDLRLKLCLEYCFSTDLRTSKSLRAHVRSKYKLF  
 KAYKRPSSTSSQVETLEEDIPAGYGFYFVSQVSGSGSELDIYGEPLDMA  
 APRFNVLAYWKDNCRCFKLSMACDVLSPITTVASESFSIGSGVLKYRSSLIP  
 ENICALLCTRNLWLRGFPKEGEEVEEKEEKEEKEA"  
 6527..10142  
 /gene="F13F21.2"  
 join(6527..6667,6772..6859,7159..7311,7597..7757,  
 7840..8514,8612..8939,9020..9240,9369..9546,9658..10142)  
 /gene="F13F21.2"  
 /note="Hypothetical Protein"  
 /codon\_start=1  
 /protein\_id="AAD43147.1"  
 /db\_xref="GI:5430747"  
 /translation="MSENKVKRQVFAGCNCVRVNVNNSGASGLVSFGAONAVAVF  
 CPKTAQILITLPGHKASVNCVTKLPTSKFAKAKLQYLLSGSDSGIILWELSTL  
 NNDVKFPYSRSDLSPFAVSLSLIPKCKVCLDSICVDTKAIVTLSLAELPQNPGRF

gene  
 CDS  
 gene  
 CDS

```
ALALGGLDNKIKLISGERTGKTSVCELKGTHTDWRSLRSLDFSLPLHTTETIPNSIMLVS
SSQDKVIRIKVLKLVGDVSWREITLASIIEGPFVFGTFTYQISVESVLIGHEDWY
YSVEQPPVIDIGRLVNHQPLSLISASMDTMMIWRPEKTKGVNVNVVCGELSHC
ALGFCVGHSPNSLSLAHGYGAGFLMRNVSSSEKENWQKVPSPGSHGFAAVTDTWT
ARTGEVLLSODOTTRVFSANKNDENAEDEHWEHLARPQVHGHDINCVMVCGKG
NHRFVSGAEKVVVEAPLPLKTLNHTCAGGEGSPEDQADQVQLGANMSALGLS
QKPIVLHSSSEPLERNGGEGDITETVEAPAEKPEPPIEDQAFHTLPFESHKLY
GHNEFLSCLSDHKNLVASSCAQASNAEILWMEVGTWKAIVGRQLSHSLTVTHLEF
SYDDTLLSVDRHFSFVSIGORTDNGEVSHKMAVEAKRIIWAQSNWPFQHPAT
SDDTKTKIWSVENDIRIKIQLVLPFGSSVTAVAWTGLDRNEKSGCVAVGMESGLIE
LGNKVIETEEGTATAALALRELFPMCHVSANRLWRPTEKCESNQSLRWLTSCGD
DNCVRFNEFKF
complement (10370..11380)
/gene="F13F21.3"
/codon_start=1
/gene="F13F21.3"
/note="97% identical to geranyl pyrophosphate
synthase [Arabidopsis thaliana] gi|2578822."
/product="geranyl geranyl pyrophosphate synthase"
/protein_id="AAD43148.1"
/db_xref="GI:5430748"
/translation="MRPRYSLLSAMLRLPSNRRLSSIASDSEFISVMKNKAKSIN
KALDNSIPLCNNFVLEWPLVSVHKAMRYTLPGCKVRPMLCLVACELVGGQESTAM
PAACAVEMIAASLIIDDLPCWDDSLRKGKFTNHKVFGKTSIIASNALRSLAVKQT
LASTSLGVTSEVRVAVQEMARAVGTGLVAGQADLAGERMSFKNEDDELRYLELMH
VHKTAVLVEAAAVGAIMGGSDSEIERLKSVCARCVGLMFQVMDVLDLTKSSEELGK
TAGKDLITGKLTYPKVMGVNDAREYAKRLNREAQHLQGFSDSKVPLLSLADYIVKR
QN"
complement (12218..14173)
/gene="F13F21.4"
complement (join(12218..12318,12359..12418,12518..12695,
12784..12898,12987..13143,13236..13410,13500..13602,
13902..14173))
/gene="F13F21.4"
/note="Hypothetical Protein"
/codon_start=1
/protein_id="AAD43149.1"
/db_xref="GI:5430749"
/translation="WVSDSLVTLQREILRSLDLETTTPASVRRQLEVVFGVELTDKK
AFVRQIDAFESLDALLESKPEEDCNGQONDESGENDDDKTELPAKAKGGGF
NKICQLSPLEKFLGTSQARTVEVVKWYIREHDQDPNTRNILLDESILHSLFRV
KTNFMQNKALAKHIALWNLDDGCFNKVEEDVDETSERDEKDVKEEALHNEEE
SREEDFQSKRRKRPKAKSEERKPKGGGTVCVSLSPQLQFTGPQLARTFV
KMLWKYIKENNLQPSDKRTIICDESLSLFPVESINFMQNKOLAKHIWPLVORDEA
GTNDEPKGQKMKMETDEVKVNIKSSTFLNYIDNESNEEKATSSRIKTEE"
complement (15762..16883)
/gene="F13F21.5"
complement (join(15762..15977,16061..16492,16809..16883))
/note="Unknown Protein; Location of ests VBVDQ12
(gi|757594 and gi|757593)"
/codon_start=1
/protein_id="AAD43150.1"
/db_xref="GI:5430750"
/translation="MAMSHLFLSSSPQSGSLALRLHSTTQFTLSYKNNKNCDSFQANE
AKVSRSLICRAIHMHSGSGQPKPLNFDNLLRTHKWDNSPOPVKEFPWNRAPGN
FTQLVDLAISSVVKLFVPIPLAVSSISEMSYCAHERKALVPFPLVIGWVAGVLOET
ALKISPLKEASVPHLITAMMFFTLIKLPGYPYWGRLVLPHPFANGVLLRALWSMF
FWYKTRNTSGNPLQNHSLETE"
21730..21921
/gene="F13F21.6"
21730..21921
/gene="F13F21.6"
/note="Unknown protein; Location of ests 164G1817
(gi|2764114), 114G10T7 (gi|2597578), and 114G117
(gi|2597579)"
/codon_start=1
/protein_id="AAD43151.1"
/db_xref="GI:5430751"
/translation="WVTLKWEICIELLKLTVDVFAVAESIEVAFRHRPPVVIQSSV
MYGRSNTYAVPIPLVGLF"
23803..26346
```

```
/gene="F13F21.7"
23803..26346
/gene="F13F21.7"
/note="Hypothetical Protein"
/codon_start=1
/protein_id="AAD43152.1"
/db_xref="GI:5430752"
/translation="MERPFQGFILLISYTVVATFDDPPSPENADITKLEQKCF
INSDVNLKFNDRLKRAYIALQWGLKAIYSDPFKTTANWGSVDYGVYCAPALD
DKSLTVAGVADLNHADIAGHLPELGLMTDLALPHINSRFGCIIPKSLSKLALVFEF
DVNNRFGQPEVLSWPSLKFOLDLRYNEPEGLSPRIFDKDLDAIPLNNRFSVI
PGTIGKASVUTTEANNKFSKIPKISGNMKNLIEIVTGNLTCFNEIGLLNVT
VFDASKGVSLPSTLSGLASBOLDLSHNKLTVVDKFCMLNLSFKFSINFFN
GEASCVPGRNKGQFDDTNCLQNRSPQKAKQCLPVSVREDCSKQCSGNGSG
SPSPNPTSPKPSKPEFVMPKPSDSKPTETPKTPEQSPKQPPKESPKPEPEN
KHELKQXESKQPSKPEFVMPKPSDSKPTETPKTPEQSPKQPPKESPKPEPEN
DDVDASPVKNRRSPKPKVEDTRVPQPPMPSKPSPIYSPVPVHSPPPVYSS
PPPHVYSPPPVSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
PPPVYSPPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
PVQAPTQVQSSPTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
SSEGNQSQAPTFILEPHAPTNSKPVQSPSPSPSPSPSPSPSPSPSPSPSPSPSP
SSVSSSPSTSTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
28449..29313
/gene="F13F21.8"
join(28449..28680,28766..28965,29065..29313)
/gene="F13F21.8"
/note="Unknown Protein"
/codon_start=1
/protein_id="AAD43153.1"
/db_xref="GI:5430753"
Alignment Scores:
Pred. No.: 4,21e-79 Length: 125021
Score: 1098.00 Matches: 244
Percent Similarity: 70.18% Conservative: 22
Best Local Similarity: 64.38% Mismatches: 25
Query Match: 61.14% Indels: 89
DB: 8 Gaps: 11
US-10-088-187a-11 (1-341) x AC007504 (1-125021)
QY 33 GluLeuSerValAlaValAlaLeuThrValProAspGlyHisValTrpArgValGlyLeu 52
Db 28233 GAGCTTTTCAGTTGTTTTCGACATCAACATCCTGATGTCATGTTTG-CATGTAGACTA 28291
QY 53 ArgLysAlaAspAsnLysIleTrpPheGlnAspGlyTrpGlnGluPheValAspArgTyr 72
Db 28292 AGGAAAGCCAAACAAATCTGGTTTCAAGACGGTTGCGAGAGTTTGTCAACCGTTTC 28351
QY 73 SerIleArgIleGlyTyrLeuLeuIlePheArgTyrGluGlyAsnSerAlaPheSerVal 92
Db 28352 TCCATTCGATGTTGTT-----TTGATATACAAA-----GTTACAGTC 28387
QY 93 TyrIlePheAsnLeu---Ser-HisSerGluIleAsnTyrHisSerThrGly----- 108
Db 28388 TACATTTTCAATTTTATCTCCACACTTCAGATCAACACCACTTCTAGTAGTGAAGCTCT 28447
QY 109 ----LeuMetAspSerAlaHisAsnHisPhe---LysArgAlaArgLeuPheGluAspLe 126
Db 28448 TATGCAATGATGATCCCGCAGCAATCAGTTTCAACAAACGTCGCTCGATGTTTGAAGATCC 28507
QY 126 uGluAspGluAspAlaGluValIlePheProSerSerValTyrProSerProLeuProGlu 146
Db 28508 TGAACCTCAAGATGCTAAGGTCATTTATCCATCCGAAAC-----CCTCA 28549
QY 146 uSerThrValProAlaAsnLysGlyTyr---AlaSerSerAlaIleGlnThrLeuPheTh 165
Db 28550 ATCTACTGAACCCAGTGAATAAAGGTTATGCGCGGTTTCTACAGCATCAACCAAGCTTTTCAA 28609
QY 165 rGlyProValLysAlaGluGluProThrProThrProLysIleProLysIleArgGlyAr 185
Db 28610 A---GAATCTAAAGCTGAAGAA-----ACGCCCAAGGACTTAAAGAGAGAGGAAG 28657
```



JOURNAL regions of the genome  
DNA Res. 10 (1), 27-33 (2003)  
MEDLINE 22579290  
PUBMED 12693552  
REFERENCE 2 (bases 1 to 78589)  
Sato, S.  
Direct Submission

Submitted (12-DEC-2002) Shusei Sato, Kazusa DNA Research Institute,  
Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,  
Chiba 292-0818, Japan (E-mail: ssato@kazusa.or.jp,  
URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935(ex.2337),  
Fax: 81-438-52-3934)

## FEATURES

## source

Location/Qualifiers  
1..78589  
/organism="Lotus corniculatus var. japonicus"  
/mol\_type="genomic DNA"  
/variety="japonicus"  
/db\_xref="taxon:34305"  
/chromosome="2"  
/clone="LjT03H18"  
/clone\_lib="LjT library"  
/note="TAC clone: TW0262-synonym: Lotus japonicus"

## ORIGIN

## Alignment Scores:

Pred. No.: 8,23e-59 Length: 78589  
Score: 850.00 Matches: 195  
Percent Similarity: 48.79% Conservative: 46  
Best Local Similarity: 39.47% Mismatches: 71  
Query Match: 47.33% Indels: 182  
DB: 8 Gaps: 11

US-10-088-187A-11 (1-341) x AP006145 (1-78589)

QY 22 ValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeuThr 41  
DB 16037 ATTCCGGACAAATTTCTAGGAAATATGCGGTACAGCTTCACAATTCCTACCTCACT 16096  
QY 42 ValProAspGlyHisValTyrArgValGlyLeuArgLysAlaAspAsnLysIleTyrPhe 61  
DB 16097 GTTCCAGATGGTACTGTTGGGTGTAGGGTTAAGAGCGCCGACACAGAAATTTGGTTT 16156  
QY 62 GluAspGlyTyrGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 81  
DB 16157 GCTGATGGTGGCAAGAAATTTCTCAACGCTACCATGGCATTTGGATCTTTTAGTA 16216  
QY 82 PheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSerGlu 101  
DB 16217 TTTCAATGTAAGGGAAATTCATCTTTTCATTTTCATATTTTCAATATGAGCACTTCTGAG 16276  
QY 102 IleAsnTyrHisSerThr-----GlyLeuMetAspSerAlaHisAsn 115  
DB 16277 GTGAATATCAGTCTGCACTGAGAGCCGTAATGAAGGCCCTGCTTTACGAATATCAT 16336  
QY 116 HisPheLysArgAlaArgLeuGluAspGluLeuAspGluAspAlaGluValIlePhe 135  
DB 16337 CAT-----ATTTTTCGAAATGGAAGATACAGACTCCTTCGAATTCATG 16381  
QY 136 ProSerSerValTyrProSerProLeuProGluSerThrValProAlaAsnLysGlyTyr 155  
DB 16382 GATTCCTCT-----CCTTCAAAACCTTACTCTCTGGTTTCATG-----CAAAACCAAGGTTTTT 16432  
QY 156 AlaSerSer-----AlaIleGlnThrLeu 163  
DB 16433 GTGGATCTGTAGATCACTGACCTGGGAAGAAATCAATCTCCTGACGGCAGAAATTTG 16492  
QY 164 PheThrGlyProValLys-----AlaGluGluProThrProThrPro 177  
DB 16493 TTTAATGGTGGTCTAAACTCAATCGTGTGAAGTGGGGAGATGAGTGGAGTCTTTCTCT 16552  
QY 177 ----- 177  
DB 16553 TCAAAGGATGCCAAATTCACAGGGCAATCACTCAACAGAGATATAGGTGTTTCAGTTAAT 16612

QY 178 -----LysIlePro----- 180  
DB 16613 GCTGTTGAATTTAAAGGTCTACTCAAGAATTTGAAATTCGTTATCTCTAACAATGAGGAA 16672  
QY 181 -----LysLysArgGlyArgLysLysAsnAlaAspProGluGluLeu----- 195  
DB 16673 GGGGTTAGCAAACTGTGAAAGGAAGAGGAAGTCTGAACCAAGTAAGTTCTTTTAGCCA 16732  
QY 195 ----- 195  
DB 16733 AGTTTACTTTTATGATCTGTAATCGAAATAATGAATTTTCTCATCAATGTAATTGCA 16792  
QY 196 -----AsnSerSerAlaProArgAspAspAspProGluAsnArgSer 209  
DB 16793 GAACTGTAGATGGTCAGGAACCTTCTGCTGAAATGAAGAGGAGGACAGAAATGCGCTAT 16852  
QY 210 LysPheTyrGluSerAlaSerAlaArgLysArgThrValThrAlaGluGluArgGluArg 229  
DB 16853 AGATTTTATGAAAGTCTTCTGTAGAAAGAGAACCGTACTGCGAGGAGAAAGAGAGG 16912  
QY 230 AlaIleAsnAlaAlaLysThrPheGluProThrAsnProPheArgValValLeuArg 249  
DB 16913 GCATTAATGATGCGAAAGCATTTGAACAGCTAATCTTTCTGTCAGTTGCTCTTCA 16972  
QY 250 ProSerTyrLeuTyrArgGlyCysIleMet----- 259  
DB 16973 CCTCTCTATTTATATAGGGATGCATATGTTGAGTTGTTGATTTGATATTTAT 17032  
QY 259 ----- 259  
DB 17033 CTTTATTCTTCTCAACATCAGGGTTTAAATTTGCGATCCACAGAGCAATTTGCGCCGC 17092  
QY 259 ----- 259  
DB 17093 AACATAACGGTTTTCAGTCCCAACGGCATCAACATGCAATTTGTCGGCTTCAATCT 17152  
QY 259 ----- 259  
DB 17153 GCATTTGTTGCGCAATTCGCAATGTGACCGCACTGCAATTTAAACCTGTTCAACATG 17212  
QY 260 -----TyrLeuProSerGlyPheAlaGluLys 268  
DB 17213 TCATCTTCATGCTAACACCTTCTGTTTCCAGTATCTGCTTCTGCTTTGCGAGAAAAG 17272  
QY 269 TyrLeuSerGlyIleSerGlyPheIleLysValGlnLeuAlaGlu-----LysGlnTyrPro 287  
DB 17273 AATTTGAATGGGTTTTCGGATTCATCAAACTTCAGTTGCGGAGAGAGAGAGTGGCCA 17332  
QY 288 ValArgCysLeuTyrLysAlaGlyArgAlaLysPheSerGlnGlyTyrTyrGluPheThr 307  
DB 17333 GTTCGCTGCTATATAGAGAGGTGAGCAAAAGTTAAGCCAGGGTTGTTGTAATTCACC 17392  
QY 308 LeuGluAsnAsnLeuGlyGluGlyAspValCysValPheGluLeuLeuArgThrArgAsp 327  
DB 17393 TTAGAGAACAAATTTAGGGGAAGGGATGCTGTGTGTTTGTAGCTGCTCAGAACGAGGAT 17452  
QY 328 PheValLeuLysValThrAlaPheArgValAsnGluTyrVal 341  
DB 17453 GTAGTGTGCAAGTACTTTTATTTTCTGTTAACTGAAGAAGTG 17494  
RESULT 15  
AC137825/c 243777 bp DNA linear HTG 13-AUG-2004  
LOCUS Medicago truncatula clone mth2-23j1, WORKING DRAFT SEQUENCE, 17  
DEFINITION unordered pieces.  
AC137825  
VERSION AC137825.22 GI:50897368  
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.  
SOURCE Medicago truncatula (barrel medic)  
ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 30, 2004, 06:11:13 ; Search time 570 Seconds

(without alignments)  
3140.443 Million cell updates/sec

Title: US-10-088-187A-11

Perfect score: 1796

Sequence: 1 MPRPFHKLIFSSTIQEKL.....LRTDRFLVKVTAFRNVEYV 341

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlh  
-Q/cn2\_1/USPTO spool/US1008187/runat\_29122004.115842.107/app\_query.fasta\_1.519  
-DB=N\_Geneseq\_23Sep04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US1008187 @CN 1.1 470 @RUNAT 29122004.115842.107 -NCPUS=6 -ICPU=3  
-NO\_WRAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_23Sep04:\*

1: Geneseqm1980s:\*

2: Geneseqm1990s:\*

3: Geneseqm2000s:\*

4: Geneseqm2001as:\*

5: Geneseqm2001bs:\*

6: Geneseqm2002as:\*

7: Geneseqm2002bs:\*

8: Geneseqm2003as:\*

9: Geneseqm2003bs:\*

10: Geneseqm2003cs:\*

11: Geneseqm2003ds:\*

12: Geneseqm2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1796	100.0	1026	12	Adn73922 Thale cre
2	1796	100.0	1495	4	Aaf62446 A thalian
3	1789	99.6	1510	3	Aac39629 Arabidops
4	1783	99.3	1509	3	Aac51591 Arabidops
5	1711.5	95.3	1347	3	Aac51587 Arabidops
6	1601.5	89.2	5000	4	Aaf62416 A thalian

RESULT 1

ADN73922

ID ADN73922 standard; cDNA; 1026 BP.

XX

AC ADN73922;

XX

DT 15-JUL-2004 (first entry)

XX

DE Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 1817.

XX

KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;

KW growth regulator; animal feed product; thale cress;

KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.

XX Arabidopsis thaliana.

XX

FN WO2004035798-A2.

XX

PD 29-APR-2004.

XX

PF 20-OCT-2003; 2003WO-EP011658.

XX

PR 18-OCT-2002; 2002EP-00079408.

XX (CROP-) CROPDESIGN NV.

PA

XX Inze D, De Veylder L, Vlieghe K;

PI

XX WPI; 2004-348466/32.

DR P-PSDB; ADN73923.

DR

ALIGNMENTS

Aac54934 Arabidops  
Aac3657 Arabidops  
Aac37617 Arabidops  
Aac37903 Arabidops  
Abz12468 Arabidops  
Abx2122 Arabidops  
Adc03457 Wheat flo  
Aac3415 Rice flo  
Aac35159 Arabidops  
Abz15168 Arabidops  
Aac51592 Arabidops  
Aac51585 Arabidops  
Aac49013 Arabidops  
Aac36162 Arabidops  
Aac33396 Arabidops  
Aac49003 Arabidops  
Abz12507 Arabidops  
Aq92625 Cotton ex  
Adc03363 Rice flo  
Aac51560 Arabidops  
Aac33640 Arabidops  
Abx2454 Arabidops  
Aac03458 Wheat flo  
Ab173975 Corn tass  
Abz14722 Arabidops  
Aac54740 Arabidops  
Aac34223 Arabidops  
Aac54739 Arabidops  
Aac33412 Arabidops  
Aac54709 Arabidops  
Aac54731 Arabidops  
Ada4821 Rice gene  
Adj48254 Maize oil  
Aac3604 Arabidops  
Aac3404 Arabidops  
Aac3321 Chemical  
Abq29902 Oligonucl  
Abq29903 Oligonucl  
Aac54411 Arabidops

XX Altering plant characteristics, useful for producing plants for enzyme or  
PT pharmaceutical production comprises modifying in a plant, expression of  
PT one or more nucleic acids and/or modifying level or activity of one or  
PT more proteins.

XX Claim 1; SEQ ID NO 1817; 134pp; English.

XX This invention relates to a novel method for altering one or more plant  
CC characteristics. Specifically, it refers to identifying genes that are up  
CC - or down-regulated in transgenic plants overexpressing the heterodimeric  
CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to  
CC alter plant characteristics accordingly. The present invention describes  
CC generating transgenic plants for the production of growth regulators,  
CC enzymes, therapeutic, pharmaceuticals and animal feed products, where  
CC the altered plant characteristics are selected from increased yield or  
CC biomass, enhanced survival capacity, stress tolerance, plant architecture  
CC or physiology, altered endoreduplication, biochemistry, signal  
CC transduction, storage lipid mobilisation and/or altered photosynthesis,  
CC each relative to the corresponding wild type plants. Accordingly, these  
CC sequences can also be useful as positive or negative selectable markers  
CC during transformation of cells or tissues. The identified genes play a  
CC role in a variety of biological processes such as DNA replication, cell  
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as  
CC transcription factors. This polynucleotide sequence is thale cress CDNA  
CC repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa  
CC transcription factor, given in an exemplification of the invention.

XX SQ Sequence 1026 BP; 295 A; 235 C; 228 G; 268 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.57e-176 Length: 1026  
Score: 1796.00 Matches: 341  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-10-088-187A-11 (1-341) x ADN73922 (1-1026)

QY 1 MetProArgProPheHisLysLeuIlePheSerSerThrIleGlnLysArgLeu 20  
DB 1 ATGCCAGCCCTTCTTCCATAGTTGATTTCTCATCTCATCTATCCAGAAAAAGCTCTG 60  
QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40  
DB 61 AGGGTCCCAAGATAAGTTGTGAGTAAATTCAGAGTAAATTCAGAGTAAATTCAGAGTAA 120  
QY 41 ThrValProAspGlyHisValTtpArgValGlyLeuArgLysAlaAspAsnLysIleTtp 60  
DB 121 ACAGTACCTGATGCTATGTTGGCGGTGAGGACTAAGGAAAGCTGACACAAATTTGG 180  
QY 61 PheGlnAspGlyTtpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80  
DB 181 TTTCAAGATGTTGGCAAGAGTTGTTGACCGTTACTCCATTCGCAATGCTTATCTTTTG 240  
QY 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100  
DB 241 ATTTTATAGATATGAAGAAACTCGCTTCAGCGTCTACATTTTCAATTTATCCCACTCT 300  
QY 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120  
DB 301 GAGATCATTTACATTTCCACCGTCTCATGATTCGCTCACAACCACTTCAACCGGCC 360  
QY 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr 140  
DB 361 CGTTTGTGTTGAAGACCTTGAAGATGAAGATGCGGAGTCACTTCTCTCTCTCTCTGTG 420  
QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160  
DB 421 CCATCACCATCTCTCTGAGTCTACAGTACCAAGCAACAAAGGGTATGCTAGTTCAGGCATC 480  
QY 161 GlnThrLeuPheThrGlyProValLysAlaGluLeuProThrProThrProLysIlePro 180

DB 481 CAAACCTTGTCTACTGACAGTTAAAGCTGAAGAGCAACCAACCCCAAAATACCT 540  
QY 181 LysLysArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaPro 200  
DB 541 AAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
QY 201 ArgAspAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220  
DB 601 CGAGATGATGATCCAGAGAACCGTTTCAAGTTCTACAGAGTGTCTTCTCGAGAGAGAGA 660  
QY 221 ThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThr 240  
DB 661 ACCGTGACTCAG 720  
QY 241 AsnProPhePheArgValValLeuArgProSerTyrIleTyrArgGlyCysIleMetTyr 260  
DB 721 AACCTTTCTTCAGAGTGTCTTCTGCGACCATCTATCTATACAGAGAGTTCATCATGTAT 780  
QY 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyLysSerGlyPheIleLysValGln 280  
DB 781 CTTCCTTCTGCGTGTCTGAGAGTACTTAAGTGGGATCTCCGGTTCATCAAGTCCAG 840  
QY 281 LeuAlaGluLysGlnTtpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSer 300  
DB 841 CTTGCGGAGAGAAACAATGCGCTGTCGATGTCTCTACAAAGCGGAGAGGAGGAGGAGGAG 900  
QY 301 GlnGlyTtpTyrGluPheThrLeuGluAsnAsnLeuGlyGluGlyAspValCysValPhe 320  
DB 901 CAAGGATGGTACGAATTCATCTAGAGAACCACTTAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
QY 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyr 340  
DB 961 GAGCTGCTCAGAACCAAGAGATTGTTGTTTGAAGTGAAGCTTTCGAGTCAACGAGTAC 1020  
QY 341 Val 341  
DB 1021 GTC 1023

RESULT 2

AAF62446  
ID AAF62446 standard; cDNA; 1495 BP.  
XX AAF62446;  
XX 05-NOV-2001 (first entry)  
XX A thaliana VRN1 coding sequence.  
XX VRN1; vernalisation; flowering; crop; ss.  
XX Arabidopsis thaliana.  
XX Key Location/Qualifiers  
CDS 269..1294  
FT /\*tag= a  
FT /product= "VRN1"  
XX WO200121822-A1.  
XX 29-MAR-2001.  
XX 13-SEP-2000; 2000WO-GB003525.  
XX 17-SEP-1999; 99GB-00022071.  
XX (PLAN-) PLANT BIOSCIENCE LTD.  
XX Dean C, Levy YI;  
PI WPI; 2001-273467/28.  
DR P-PSDB; AAB35491.  
XX

PT Novel VRN1 polynucleotide sequence encoding a polypeptide which alters  
 PT vernalization response of plant in which VRN1 nucleic acid is expressed,  
 PT useful for influencing and assessing vernalization phenotype of plants.  
 XX  
 PS  
 XX Claim 2; Fig 7; 91pp; English.

CC The present invention provides the protein and coding sequences of  
 CC Arabidopsis thaliana VRN1. This protein is capable of altering the  
 CC vernalisation responses of a plant. Also provided are a number of PCR  
 CC primers used to isolate the sequences. The sequences are useful in the  
 CC production of crop plants, where they are able to control the timing of  
 CC flowering, the duration of vernalisation required, the optimum  
 CC temperature, or even eliminate the need for vernalisation completely. The  
 CC present sequence is the VRN1 coding sequence

XX  
 SQ Sequence 1495 BP; 403 A; 322 C; 312 G; 458 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2,698-176 Length: 1495  
 Score: 1796.00 Matches: 341  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-088-187A-11 (1-341) x AAF62446 (1-1495)

QY 1 MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20  
 DB 269 ATGCCACGGCCCTTCTTCATAGTGTGATTTCTCATCCACTATCCAGAAAAGCTGTG 328  
 QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40  
 DB 329 AGGGTCCCAGATAAGTTTGTGAGTAAATTCAGAGGATGAGCTTCGGTGTGTCACATC 388  
 QY 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60  
 DB 389 ACAGTACCTCGTGGTCATGTTTGGCGGTGAGGACTAAGGAAAGCTCACAAACAAATTTGG 448  
 QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80  
 DB 449 TTTCAAGATGGTTGGCAAGAGTTTGTGACCGTTACTCCATTCGGATGTTATCTTTTG 508  
 QY 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100  
 DB 509 ATTTTATAGATAGAGGAACCTCGCTTCAGCGTCTACATTTTCAATTTATCCCACTCT 568  
 QY 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120  
 DB 569 GAGATCAATTACCATTTCCACCGGTCTCATGGATTCGGCTCACACCACTTCAACGCGCC 628  
 QY 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr 140  
 DB 629 CGTTTGTGTAAGACCTTGAAGATGAAGTCCGAGGTCACTTCTCTTCTGTGTGATC 688  
 QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160  
 DB 689 CCATCACCACTTCTCGAGTCTACAGTACCAGCCCAACAAAGGTATGCTAGTTCCAGCCATC 748  
 QY 161 GlnThrLeuPheThrGlyProValLysAlaGluLuproThrProThrProLysIlePro 180  
 DB 749 CAACCTTGTCTCACTGGACCACTTAAGCTGAAGAGCCCAACGCCCAACCCCAAAAATACCT 808  
 QY 181 LysLysArgGlyArgLysLysLysAsnAlaAspProGluLysLeuAsnSerSerAlaPro 200  
 DB 809 AAAAAGAGGGAGGAGGAAGAAATGCTGATCTCTGAGAAATAACTCATCAGCTCCG 868  
 QY 201 ArgAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220  
 DB 869 CGAGATGATGATCCAGAGAACCGTTCAAGTTCTACGAGGTGCTTCTCGGAGAAAGAGA 928  
 QY 221 ThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThr 240

DB 929 ACCGTGACTGCAGAGAAAGAGAGAGAGCCATCAATGCAGCCAAACGTTCCGAACAACA 988  
 QY 241 AenProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260  
 DB 989 AACCTTTCTTCAGAGTGGTTCTGGACCATCTCTATCTATACAGAGGTTGCATCATGTAT 1048  
 QY 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGln 280  
 DB 1049 CTTCTCTCTGGGTTTGTCTGAGAAGTACCTAAGTGGGATCTCCGGGTTTCATCAAGTCCAG 1108  
 QY 281 LeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSer 300  
 DB 1109 CTTGCGGAGAAACAATGGCCCTGTTCTGATGTCTCTACAAAGCCGGAGAGCCAAATTCAGT 1168  
 QY 301 GlnGlyTyrTyrGluPheThrLeuGluAsnLeuGlyGluGlyAspValCysValPhe 320  
 DB 1169 CAAGGATGGTACGAATTCACCTCTAGAGAACACTTAGGAGAGGAGACGCTCTGTGTGTTT 1228  
 QY 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValLeuGluTyr 340  
 DB 1229 GAGCTGCTCTCAGAACCAAGAGATTTCCGTTTGAAGTGACAGCCCTTTCGAGTCAACGAGTAC 1288  
 QY 341 Val 341  
 DB 1289 GTC 1291  
 RESULT 3  
 AAC39629  
 ID AAC39629 standard; DNA; 1510 BP.  
 XX  
 AC AAC39629;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 25332.  
 XX  
 KW Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway; metabolic pathway;  
 KW promoter; termination sequence; ss.  
 XX Arabidopsis thaliana.  
 XX  
 FN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.

Mon Jan 3 13:13:30 2005

us-10-088-187a-11.rng

PR	14-MAY-1999;	99US-0134221P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145918P.
PR	18-MAY-1999;	99US-0134768P.	PR	28-JUL-1999;	99US-0145919P.
PR	19-MAY-1999;	99US-0134941P.	PR	28-JUL-1999;	99US-0145951P.
PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999;	99US-0135623P.	PR	02-AUG-1999;	99US-0146389P.
PR	25-MAY-1999;	99US-0136021P.	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	99US-0136392P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-0136782P.	PR	04-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	99US-0137528P.	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139119P.	PR	10-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	16-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	17-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	18-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139460P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139462P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139463P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	21-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	22-JUN-1999;	99US-0140353P.	PR	30-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	31-AUG-1999;	99US-0151930P.
PR	23-JUN-1999;	99US-0140895P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140823P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140991P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-01411287P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141842P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0142154P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142390P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142803P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142920P.	PR	24-SEP-1999;	99US-0155559P.
PR	09-JUL-1999;	99US-0142977P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0143542P.	PR	29-SEP-1999;	99US-0156596P.
PR	14-JUL-1999;	99US-0143624P.	PR	04-OCT-1999;	99US-0157117P.
PR	15-JUL-1999;	99US-0144005P.	PR	05-OCT-1999;	99US-015753P.
PR	16-JUL-1999;	99US-0144085P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144086P.	PR	07-OCT-1999;	99US-0158029P.
PR	19-JUL-1999;	99US-0144325P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144331P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144334P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159329P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159637P.
PR	21-JUL-1999;	99US-0144814P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0145086P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	23-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.

```
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 1,466-175 Length: 1510
Score: 1789.00 Matches: 340
Percent Similarity: 99.71% Conservatives: 0
Best Local Similarity: 99.71% Mismatches: 1
Query Match: 99.61% Indels: 0
DB: 3 Gaps: 0

US-10-088-187A-11 (1-341) x AAC39629 (1-1510)

QY 1 MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20
286 ATGCCACGCCCTTCTTCCATAAGTTGATTTTCTCATCCACTATCCAGAAAAACGCTCG 345
QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
346 AGGGTCCCAAGATAAGTTTGTGAGTAAATTCAGAGATGAGCTTTCGGTTGCTGTGCACCTC 405
QY 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAsnLysIleTrp 60
406 ACAGTACCTGATGTCATGTTTGGGTGAGGACTAAGAAAGCTGACACAAAAATTGG 465
QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80
466 TTTCAAGATGTTGGCAAGAGTTTGTGACCGTTACTCCATTCGCATGATTATCTTTG 525
QY 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100
526 ATTTTATAGATATGAAGAACTCTGCTTTCAGCGTCTACATTTTCAATTTATCCACTCT 585
QY 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120
586 GAGATCAATTACATTCACCGGTCTCATGATTCGCTCAACACCTTCAACACGCGCC 645
QY 121 ArgLeuPheGluAspLeuGluAspAlaGluValIlePheProSerSerValTyr 140
646 CGTTTGTGTTGAAGACCTTGAAGATGAGATGCGGAGGTATCTTCTTCTCTGTGTAC 705
QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160
706 CCATCACCACCTTCTCTGAGTCTACAGTACCAGCCCAACAAAGGTATGCTAGTTTCAGCCATC 765
QY 161 GlnThrLeuPheThrGlyProValLysAlaGluProThrProThrProLysIlePro 180
766 CAATCCTTGTTCACCTGGACAGTAAAGCTGAAGAGCCCAACGCCCAACCAAAATACCT 825
QY 181 LysLysArgGlyValGlyLysLysAsnAlaAspProGluGluIleAsnSerSerAlaPro 200
826 AAAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 885
QY 201 ArgAspAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220
886 CGAGATGATGATCCAGAGAACCGTTCAAGATTTCTACGAGGTGCTTCTCGGAGAAAGAGA 945
QY 221 ThrValThrAlaGluGluArgGluArgAlaIleAsnAlaLysThrPheGluProThr 240
946 ACCGTGACTGCAAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1005
QY 241 AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260
1006 AACCTTCTTCAGAGTGTCTGCGACCATCTCTATCTATACAGAGGTGTCATCATGTAT 1065
QY 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGln 280
```

```
Db 1066 CTTCTTCTGGGTTTGCAGAAAGTACCTAAGTGGGATCTCCGGGTTTCATCAAGTCCAG 1125
QY 281 LeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSer 300
1126 CTTGGGAGAAACAATGGCTGTTTCGATGCTCTACAAAGCCGGGAGAGCCAAATTCAGT 1185
QY 301 GlnGlyTrpTyrGluPheThrLeuGluAsnLeuGluGlyAspValCysValPhe 320
1186 CAAGGATGTTACGAATTCACCTAGAGAACAACTTAGGAGAAAGGAGACGCTCTGTGTGTTT 1245
QY 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyr 340
1246 GAGCTGCTCAGAACCAAGATTTGCTTTTGAAGTGACAGCCCTTTTCGAGTCAACGAGTAC 1305
QY 341 Val 341
1306 GTC 1308

RESULT 4
AAC51591
ID AAC51591 standard; DNA; 1509 BP.
XX
AC AAC51591;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 69076.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 08-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0129845P.
PR 21-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
```

Mon Jan 3 13:13:30 2005

us-10-088-187a-11.rng

PR	25-MAY-1999;	99US-0136021P.	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	99US-0136392P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-0136782P.	PR	05-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	99US-0137528P.	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	99US-0138094P.	PR	08-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139111P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	17-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149802P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140354P.	PR	07-SEP-1999;	99US-0152363P.
PR	24-JUN-1999;	99US-0140695P.	PR	10-SEP-1999;	99US-0153070P.
PR	28-JUN-1999;	99US-0140823P.	PR	13-SEP-1999;	99US-0153758P.
PR	29-JUN-1999;	99US-0140991P.	PR	15-SEP-1999;	99US-0154018P.
PR	30-JUN-1999;	99US-0141287P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0141842P.	PR	20-SEP-1999;	99US-0154779P.
PR	01-JUL-1999;	99US-0142154P.	PR	22-SEP-1999;	99US-0155139P.
PR	02-JUL-1999;	99US-0142055P.	PR	23-SEP-1999;	99US-0155486P.
PR	06-JUL-1999;	99US-0142390P.	PR	24-SEP-1999;	99US-0155659P.
PR	08-JUL-1999;	99US-0142803P.	PR	28-SEP-1999;	99US-0156458P.
PR	09-JUL-1999;	99US-0142920P.	PR	29-SEP-1999;	99US-0156596P.
PR	12-JUL-1999;	99US-0142977P.	PR	04-OCT-1999;	99US-0157117P.
PR	13-JUL-1999;	99US-0143542P.	PR	05-OCT-1999;	99US-0157753P.
PR	14-JUL-1999;	99US-0143624P.	PR	07-OCT-1999;	99US-0158029P.
PR	15-JUL-1999;	99US-0144005P.	PR	08-OCT-1999;	99US-0158232P.
PR	16-JUL-1999;	99US-0144086P.	PR	12-OCT-1999;	99US-0158369P.
PR	16-JUL-1999;	99US-0144085P.	PR	13-OCT-1999;	99US-0158294P.
PR	19-JUL-1999;	99US-0144325P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144331P.	PR	14-OCT-1999;	99US-0159330P.
PR	19-JUL-1999;	99US-0144332P.	PR	14-OCT-1999;	99US-0159331P.
PR	19-JUL-1999;	99US-0144333P.	PR	14-OCT-1999;	99US-0159637P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159638P.
PR	19-JUL-1999;	99US-0144335P.	PR	18-OCT-1999;	99US-0159584P.
PR	20-JUL-1999;	99US-0144352P.	PR	21-OCT-1999;	99US-0160741P.
PR	20-JUL-1999;	99US-0144632P.	PR	21-OCT-1999;	99US-0160767P.
PR	20-JUL-1999;	99US-0144884P.	PR	21-OCT-1999;	99US-0160770P.
PR	21-JUL-1999;	99US-0144814P.	PR	21-OCT-1999;	99US-0160814P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160815P.
PR	21-JUL-1999;	99US-0145088P.	PR	22-OCT-1999;	99US-0160980P.
PR	22-JUL-1999;	99US-0145085P.	PR	22-OCT-1999;	99US-0160981P.
PR	22-JUL-1999;	99US-0145087P.	PR	22-OCT-1999;	99US-0160989P.
PR	22-JUL-1999;	99US-0145089P.	PR	25-OCT-1999;	99US-0161404P.
PR	22-JUL-1999;	99US-0145152P.	PR	25-OCT-1999;	99US-0161405P.
PR	23-JUL-1999;	99US-0145145P.	PR	25-OCT-1999;	99US-0161408P.
PR	23-JUL-1999;	99US-0145218P.	PR	26-OCT-1999;	99US-0161359P.
PR	23-JUL-1999;	99US-0145224P.	PR	26-OCT-1999;	99US-0161360P.
PR	26-JUL-1999;	99US-0145276P.	PR	26-OCT-1999;	99US-0161361P.
PR	27-JUL-1999;	99US-0145913P.	PR	28-OCT-1999;	99US-0161920P.
PR	27-JUL-1999;	99US-0145918P.	PR	28-OCT-1999;	99US-0161992P.
PR	27-JUL-1999;	99US-0145919P.			
PR	28-JUL-1999;	99US-0145911P.			
PR	02-AUG-1999;	99US-0146386P.			
PR	02-AUG-1999;	99US-0146388P.			
PR	02-AUG-1999;	99US-0146389P.			

```
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 6,12e-175 Length: 1509
Score: 1783.00 Matches: 341
Percent Similarity: 99.71% Conservatives: 0
Best Local Similarity: 99.71% Mismatches: 0
Query Match: 99.28% Indels: 1
DB: 3 Gaps: 0

US-10-088-187A-11 (1-341) x AAC51591 (1-1509)

QY 1 MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20
   |||||
Db 284 ATGCCACGCCCTTCTTCCATAAGTTGATTTCTCATCCACTATCCAGAAAAAGCTCTG 343

QY 21 -ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAla 40
   |||||
Db 344 GAGGGTCCAGATAAGTTTGTGAGTAAATTCAGGATGAGCTTCGGTTGCTGTGCACT 403

QY 40 uThrValProAspGlyHisValTrpArgValGluLysAlaAspAsnLysIleTr 60
   |||||
Db 404 CACAGTACCTGATGTGTCATGTTGGCGGTAGACTTAAGGAAGCTGACACAAATTTG 463

QY 60 pPheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeu 80
   |||||
Db 464 GTTTCAGATGTTGGCAAGAGTTGTTGACCGTTACTCCATTCGATTGGTTAICTTTT 523

QY 80 uIlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHis 100
   |||||
Db 524 GATTTTTAGATATGAAGGAACCTCTGCCCTTCAGCGCTACATTTTCAATTTATCCACTC 583

QY 100 rGluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAl 120
   |||||
Db 584 TGAGATCAATTAACCATTCACCGGCTCATGATTCGCTCACAACCACTTCAAACGGCG 643

QY 120 aArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerVal 140
   |||||
Db 644 CCGTTGTTTGAAGACCTTGAAGATGAAGATCCGCGGTCATCTTCCCTTCTCTGTGTA 703

QY 140 rProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAla 160
   |||||
Db 704 CCCATCACCACCTTCTCGAGTCTACAGTACACCAACCAAGGGTATGCTAGTTCAGCCAT 763

QY 160 eGlnThrLeuPheThrGlyProValLysAlaGluProThrProThrProLysIlePr 180
   |||||
Db 764 CCAAACTTGTCTACTGGACCAAGTTAAAGCTGAAGAGCCACCAACCCCAAAATACC 823

QY 180 oLysLysArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaPr 200
   |||||
Db 824 TAAAAGAGAGGAGGAGGAAGAAAATGCTCATCTCGAGGAATAAACTCATCAGCTCC 883

QY 200 oArgAspAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysAr 220
   |||||
Db 884 GCGAGATGATGATCCAGAACCGTTCAAAGTTCTACGAGAGTGCTTCTCGGAGAAAGAG 943

QY 220 gThrValThrAlaGluGluArgGluArgAlaIleAsnAlaLysThrPheGluProTh 240
   |||||
Db 944 AACCGTGACTGCAGAAAGAGAGAGAGCCATCAATGAGCCAAACCGTTCCGAACCAAC 1003

QY 240 rAsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTy 260
   |||||
Db 1004 AAAACCTTCTTCAGATGGTTCTCGACCACTCTATCTATACAGAGGTTCATCATGTA 1063

QY 260 rIleuProSerGlyPheAlaGluLysTyrIleuSerGlyIleSerGlyPheIleLysValGI 280
   |||||
Db 1064 TCTTCTCTCTGGTGTGTCAGAACTACCTAAGTGGATCTCCGGGTTCATCAAAAGTCCA 1123

QY 280 nLeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSe 300
   |||||
Db 1124 GCTTCGGGAGAAACAATGCCCTGTTCGATGTCTCTACAAAGCCGGAGAGCCAAATTCAG 1183
```

```
QY 300 rGlnGlyTrpTyrGluPheThrLeuGluAsnAsnLeuGlyGluGlyAspValCysValPh 320
   |||||
Db 1184 TCAGGATGGTACGAATTCACCTCTAGAGAACCACTTAGGAGAAAGGAGACGCTCTGTGTT 1243

QY 320 eGluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTy 340
   |||||
Db 1244 TGAGCTGCTCAGAACCAAGATTTTCGTTTGAAGTGAAGCAGCCTTCGAGTCAACGATGA 1303

QY 340 rVal 341
   |||||
Db 1304 CGTC 1307

RESULT 5
AAC51587
ID AAC51587 standard; DNA; 1347 BP.
XX
AC AAC51587;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 69061.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
```

us-10-088-187a-11.rng

Mon Jan 3 13:13:30 2005

PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.

PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	24-SEP-1999;	99US-0155486P.
PR	28-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Alignment Scores:	Length:	1347
Pred. No.:	Matches:	332
Score:	Conservative:	2
Percent Similarity:		84.34%

Best Local Similarity: 83.84% Mismatches: 7  
 Query Match: 95.30% Indels: 55  
 DB: 3 Gaps: 1

US-10-088-187A-11 (1-341) x AAC51587 (1-1347)

QY 1 MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20  
 |||||  
 DB 1 ATGCCACGCCCTTCTTCCTCAAGTTGATTTTCTCATCCACTATCCAGAAAAAGCTCTG 60  
 |||||  
 QY 20 ----- 20  
 DB 61 GTAAGCCGTCGTTCTTCAGCTACTGCTGTTTATGGATGTTTTGCTGAAAAATGCTCG 120  
 |||||  
 QY 20 ----- 20  
 DB 121 TTTCGATTTGATGTTCTCGCAATATTCCTGCTGTTCTTATAGATATTTGCGACATTTAT 180  
 |||||  
 QY 21 -----ArgValProAspLys 25  
 |||||  
 DB 181 ATCAATTATATGCTTCTTTATATCTCATACCGGCATGCTTTGTGCAGAGGGTCCAGATAAG 240  
 |||||  
 QY 26 PheValSerLysPheLysAspGluLeuSerValAlaValAlaLeuThrValProAspGly 45  
 |||||  
 DB 241 TTGTGAGTAAATTCAGGATGAGCTTTCGGTGTGCTGTTGCTCACTCACAGTACTGATGT 300  
 |||||  
 QY 46 HisValTTPArgValGlyLeuArgLysAlaAspAsnLysIleTTPheGlnAspGlyTTP 65  
 |||||  
 DB 301 CATGTTTGGCGTAGGACTAGGAAGCTGACACAAAATTTGGTTCAAGATGGTTGG 360  
 |||||  
 QY 66 GlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeuIlePheArgTyrGlu 85  
 |||||  
 DB 361 CAAGAGTTTGTTCACCGTTACTCCATTCGATTTGTTATCTTTGATTTTATAGATGAA 420  
 |||||  
 QY 86 GlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSerGluIleAsnTyrHis 105  
 |||||  
 DB 421 GGAAACTCTGCTTACGCTGCTACATTTTCAATTTATCCCTCTGAGATCAATTAACCAT 480  
 |||||  
 QY 106 SerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArgLeuPheGluAsp 125  
 |||||  
 DB 481 TCCACCGGTCTCATGGATTCGCTCACAACCACTTCAACGCGCGCTTTGTTGAGAC 540  
 |||||  
 QY 126 LeuGluAspGluAspAlaGluValIlePheProSerSerValTyrProSerProLeuPro 145  
 |||||  
 DB 541 CTTGAAGATGAAGATGCCGAGGTCTCTTCTCTCTCTGTTGTCATCCCATCCACTTCTCT 600  
 |||||  
 QY 146 GluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThr 165  
 |||||  
 DB 601 GAGTCTACAGTATCCAGCCCAACAAAGGTATGCTAGTTCAGCCATCAACCTTGTTCAT 660  
 |||||  
 QY 166 GlyProValLysAlaGluGluProThrProLysIleProLysLysArgGlyArg 185  
 |||||  
 DB 661 GGACCACTTAAGCTGAAGAGCCACGCCACCCCAAAATACCTTAAAGAGAGAGGAGG 720  
 |||||  
 QY 186 LysLysLysAsnAlaAspProGluLeuLeuSerSerAlaProArgAspAspAspPro 205  
 |||||  
 DB 721 AAGAAGAAAAATGCTGATCCTCAGGAATAAATCATCATGCTCCGAGATGATGATCA 780  
 |||||  
 QY 206 GluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrValThrAlaGlu 225  
 |||||  
 DB 781 GAGAACCGTTCAAGTTCTACAGAGTGTCTTCGAGAAAGAGAACCGTGTGATCAGAA 840  
 |||||  
 QY 226 GluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThrAsnProPhePheArg 245  
 |||||  
 DB 841 GAAGAGAGAGAGCCATCATCGACCAAAAGCTTCGAACCAACCAACCTTTCTTCAGA 900  
 |||||  
 QY 246 ValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyrLeuProSerGlyPhe 265  
 |||||  
 DB 901 GTGGTTCTCGCACTTCTATCTATACAGAGGTTCATCATGTATCTCTCTCTCTGGTTT 960  
 |||||  
 QY 266 AlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGlnLeuAlaGluLysGln 285  
 |||||  
 DB 961 GCTGAGAAGTACTTAAGTGGGATCTCCGGGTTTCATCAAGTCCAGTTCGCGAGAAACAA 1020  
 |||||

QY 286 TTPProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSerGlnGlyTTPTrpGlu 305  
 |||||  
 DB 1021 TGGCGTGTTCGATGTCTCTACAAAGCCGGAGAGCAAAATTCAGTCAAGGATGGTACGAA 1080  
 |||||  
 QY 306 PheThrLeuGluAsnAsnLeuGlyGlnGlyAspValCysValPheGluLeuLeuArgThr 325  
 |||||  
 DB 1081 TTCACCTCTAGAGAACACACTTAGGAGAAAGGAGACGCTGTGTGTTTGTAGCTGCTCAGAAC 1140  
 |||||  
 QY 326 ArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyrVal 341  
 |||||  
 DB 1141 AGAGATTTTCGTTTGGAAACATACAGAGGTGAAGCAGTAGATTACATA 1188  
 |||||

# RESULT 6

AAF62416  
 ID AAF62416 standard; DNA; 5000 BP.

AC AAF62416;

DT 05-NOV-2001 (first entry)

DE A thaliana VRN1 gene.

XX VRN1; vernalisation; flowering; crop; ds.

XX Arabidopsis thaliana.

XX WO200121822-A1.

PD 29-MAR-2001.

XX 13-SEP-2000; 2000WO-GB003525.

XX 17-SEP-1999; 99GB-00022071.

XX (PLAN-) PLANT BIOSCIENCE LTD.

PI Dean C, Levy YY;

DR WPI; 2001-273467/28.

PT Novel VRN1 polynucleotide sequence encoding a polypeptide which alters vernalization response of plant in which VRN1 nucleic acid is expressed, useful for influencing and assessing vernalization phenotype of plants.

PS Claim 4; Page 70-73; 91pp; English.

XX The present invention provides the protein and coding sequences of Arabidopsis thaliana VRN1. This protein is capable of altering the vernalisation responses of a plant. Also provided are a number of PCR primers used to isolate the sequences. The sequences are useful in the production of crop plants, where they are able to control the timing of flowering, the duration of vernalisation required, the optimum temperature, or even eliminate the need for vernalisation completely. The present sequence is the VRN1 gene

XX Sequence 5000 BP; 1447 A; 871 C; 931 G; 1750 T; 0 U; 1 Other;

## Alignment Scores:

Pred. No.: 2.41e-155 Length: 5000  
 Score: 1601.50 Matches: 320  
 Percent Similarity: 76.92% Conservative: 0  
 Best Local Similarity: 76.92% Mismatches: 1  
 Query Match: 89.17% Indels: 95  
 DB: 4 Gaps: 3

US-10-088-187A-11 (1-341) x AAF62416 (1-5000)

QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40  
 |||||

DB 3120 AGGGTCCCAAGATAAGTTTGTGAGTAAATTCAGGATGAGCTTTCGGTGTCTGTCACATC 3179  
 |||||

QY 41 ThrValProAspGlyHisValTTPArgValGlyLeuArgLysAlaAspAsnLysIleTTP 60  
 |||||

Db	3180	ACAGTACCTGATGGTCATGTTTGGCGTGTAGAGCTAAGGAAAGCTGACAACTTTGG	3239
Qy	61	PheGlnAspGlyTyrGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu	80
Db	3240	TTTCAAGATGGTTGGCAGAGATTGTTGACCGTTACTCCATTCGCAITGGTTATCTTTG	3299
Qy	81	IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer	100
Db	3300	ATTTTTAGATATGAAGGAACTCTGCCCTTCAGCGCTACATTTTCAATTTATCCCACTCT	3359
Qy	101	GlulIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla	120
Db	3360	GAGATCAATTACCATTCACCGCGTCTCATGGATTCGCTCACCACTTCACACGGGCC	3419
Qy	121	ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr	140
Db	3420	CGTTTGGTTGAAGACCTTGAAGATGAAGATGCCGAGGTCATCTTTCCCTTCTCTGTGTAC	3479
Qy	141	ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle	160
Db	3480	CCATCACCATTCTCTGAGTCTACAGTACCAGCCAAACAAAGGATGTAGTTACGCCATC	3539
Qy	161	GlnThrLeuPheThrGlyProValLys-----	169
Db	3540	CAAACTTGTCTCAGTGACCAAGTTAAAGTGATATTATTAACCACTGATTCCTTTATC	3599
Qy	169	-----	169
Db	3600	TATCGCTGATTACGGCTCTTATCATCTTTTTCAGGTTGATGCTTGATATTTTCCTTATCT	3659
Qy	170	----AlaGluGluProThrProThrProLysIleProLysLysArgGlyArgLysLysLys	188
Db	3660	CCAGCTGAAGAGCCAAACCCCAACCAATACCTTAAAGAGAGAGGAGGAAGAAGAA	3719
Qy	189	AsnAlaAspPro-----	192
Db	3720	ATGCTGATCCTGGTAAAGCACITTTCTCTTTGAAATGCTTCAGACTCGTTTTCAGAGGA	3779
Qy	193	-----GluGluIle	195
Db	3780	TTCAAGATCTTCTCATGATACATATATATCTTTTGATATTGTCTTACAGAGAAATA	3839
Qy	196	AsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrGluSerAla	215
Db	3840	AATCATCAGCTCCGCGAGATGATGATCCAGAACCGTTCAAAGTTCTACGAGATGCT	3899
Qy	216	SerAlaArgLysArgThrValThrAlaGluGluArgLysArgAlaIleAsnAlaLys	235
Db	3900	TCTGCGAAGAGAAACCGTGACTGCAGAAAGAAAGAGAGAGCCATCAATGCAGCCAA	3959
Qy	236	ThrPheGluProThrAsnProPheArgValValIleuArgProSerTyrLeuTyrArg	255
Db	3960	ACGTTTCGAACCAACAAACCCCTTTCTTCAGAGTGGTTCTGCGACCATCTATCTATACAA	4019
Qy	256	GlyCysIleMet-----	259
Db	4020	GGTTGCATCATGGTAATAAAAAACATCTTAGGAAGACTTAACTTTATCGGTGCTTCAC	4079
Qy	260	-----TyrLeuProSerGlyPhe	265
Db	4080	TGATCTCTAAAAGAGCCTTCTGTTTCTCTCTCAACAGTATCTTCTCTGGGTTT	4139
Qy	266	AlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGlnLeuAlaGluLysGln	285
Db	4140	GCTGAGAGTACCTTAAGTGGGATCTCCGGTTTATCAAAGTCCAGTTCGCGAGAACA	4199
Qy	286	TrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSerGlnGlyTyrTyrGlu	305
Db	4200	TGGCTGTTCGATGCTCTCAAAAGCCGGAGAGCCAAATTCAGTCAAGGATGGTACGAA	4259
Qy	306	PheThrLeuGluAsnAsnLeuGlyGluGlyAspValCysValPheGluLeuLeuArgThr	325

Db	4260	TTCACTCTAGAGAACAACTTAGGAGAAAGGAGACGTCCTGTGTGTTTGTGCTCAGAAC	4319
Qy	326	ArgAspPheValIleLysValThrAlaPheArgValAsnGluTyrVal	341
Db	4320	AGAGATTCGTTTGAAGAGTGACAGCCTTTTCGAGTCAACGAGTACGTC	4367
Db	RESULT 7		
Db	AAC54934	standard; DNA; 1394 BP.	
Qy	AAC54934;		
Db	18-OCT-2000	(first entry)	
Qy	Arabidopsis thaliana DNA fragment SEQ ID NO: 79593.		
Db	Hybridisation assay; genetic mapping; gene expression control;		
Qy	protein identification; signal transduction pathway; metabolic pathway;		
Db	promoter; termination sequence; ss.		
Qy	Arabidopsis thaliana.		
Db	EPI033405-A2.		
Qy	06-SEP-2000.		
Db	25-FEB-2000; 2000EP-00301439.		
Qy	25-FEB-1999; 99US-0121825P.		
Db	05-MAR-1999; 99US-0123180P.		
Qy	09-MAR-1999; 99US-0123548P.		
Db	23-MAR-1999; 99US-0125788P.		
Qy	25-MAR-1999; 99US-0126264P.		
Db	29-MAR-1999; 99US-0126785P.		
Qy	01-APR-1999; 99US-0127462P.		
Db	06-APR-1999; 99US-0128234P.		
Qy	08-APR-1999; 99US-0128714P.		
Db	16-APR-1999; 99US-0129845P.		
Qy	19-APR-1999; 99US-0130077P.		
Db	21-APR-1999; 99US-0130449P.		
Qy	23-APR-1999; 99US-0130510P.		
Db	28-APR-1999; 99US-0130891P.		
Qy	30-APR-1999; 99US-0131449P.		
Db	30-APR-1999; 99US-0132048P.		
Qy	30-APR-1999; 99US-0132407P.		
Db	04-MAY-1999; 99US-0132484P.		
Qy	05-MAY-1999; 99US-0132485P.		
Db	06-MAY-1999; 99US-0132486P.		
Qy	06-MAY-1999; 99US-0132487P.		
Db	07-MAY-1999; 99US-0132863P.		
Qy	11-MAY-1999; 99US-0134256P.		
Db	14-MAY-1999; 99US-0134218P.		
Qy	14-MAY-1999; 99US-0134219P.		
Db	14-MAY-1999; 99US-0134221P.		
Qy	14-MAY-1999; 99US-0134370P.		
Db	18-MAY-1999; 99US-0134768P.		
Qy	19-MAY-1999; 99US-0134941P.		
Db	20-MAY-1999; 99US-0135124P.		
Qy	21-MAY-1999; 99US-0135353P.		
Db	24-MAY-1999; 99US-0135629P.		
Qy	25-MAY-1999; 99US-0136021P.		
Db	27-MAY-1999; 99US-0136392P.		
Qy	28-MAY-1999; 99US-0136782P.		
Db	01-JUN-1999; 99US-0137222P.		
Qy	03-JUN-1999; 99US-0137528P.		
Db	04-JUN-1999; 99US-0137502P.		
Qy	07-JUN-1999; 99US-0137724P.		
Db	08-JUN-1999; 99US-0138094P.		
Qy	10-JUN-1999; 99US-0138540P.		
Db	10-JUN-1999; 99US-0138847P.		
Qy	14-JUN-1999; 99US-0139119P.		
Db	16-JUN-1999; 99US-0139452P.		
Qy	16-JUN-1999; 99US-0139453P.		

PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0158293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 20-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145919P.	PR 26-OCT-1999;	99US-0161359P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161992P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161993P.
PR 04-AUG-1999;	99US-0147204P.	PR 29-OCT-1999;	99US-0162142P.
PR 05-AUG-1999;	99US-0147302P.	Alignment Scores:	
PR 05-AUG-1999;	99US-0147192P.	Pred. No.:	2.33e-108
PR 06-AUG-1999;	99US-0147260P.	Score:	1141.50
PR 06-AUG-1999;	99US-0147303P.	Percent Similarity:	75.93%
PR 09-AUG-1999;	99US-0147416P.	Best Local Similarity:	69.63%
PR 09-AUG-1999;	99US-0147493P.	Query Match:	63.56%
PR 09-AUG-1999;	99US-0147935P.	DB:	3
PR 10-AUG-1999;	99US-0148171P.	US-10-088-187A-11 (1-341) x AAC54934 (1-1394)	
PR 11-AUG-1999;	99US-0148319P.	Length:	1394
PR 12-AUG-1999;	99US-0148341P.	Matches:	243
PR 13-AUG-1999;	99US-0148565P.	Conservative:	22
		Mismatches:	28
		Indels:	57
		Gaps:	10





Mon Jan 3 13:13:30 2005

Arabidopsis thaliana DNA fragment SEQ ID NO: 18041.  
Hybridisation assay; genetic mapping; gene expression control;  
protein identification; signal transduction pathway; metabolic pathway;  
promoter; termination sequence; ss.

DE XX  
KW KW  
KW KW  
XX OS  
XX XX  
FN EP1033405-A2.  
XX XX  
PD XX  
XX XX  
PF XX  
XX XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 08-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132456P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 19-MAY-1999; 99US-0135124P.  
PR 20-MAY-1999; 99US-0135353P.  
PR 21-MAY-1999; 99US-0135629P.  
PR 24-MAY-1999; 99US-0136021P.  
PR 25-MAY-1999; 99US-0136392P.  
PR 27-MAY-1999; 99US-0136782P.  
PR 28-MAY-1999; 99US-0137222P.  
PR 01-JUN-1999; 99US-0137528P.  
PR 03-JUN-1999; 99US-0137502P.  
PR 04-JUN-1999; 99US-0137724P.  
PR 07-JUN-1999; 99US-0138094P.  
PR 08-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 10-JUN-1999; 99US-0139119P.  
PR 14-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 16-JUN-1999; 99US-0139492P.  
PR 17-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.

41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTyr 60  
274 -----CGTCT-GGACTAAGGAAGCCAAACAACAAATCTGG 308  
61 PheGlnAspGlyTyrGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80  
309 TTTCAAGACGGTTGGCAGAGTTTGTCAACGGTTTCTCCATTCGGATTGGT----- 359  
81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeu---Ser-Hi 99  
360 ---TTCAAGATACAAA-----GTTACAGTCTACATTTTCATTTATCTCCACCA 404  
99 sSerGluIleAsnTyrHisSerThrGly-----LeuMetAspSerAlaHisAs 115  
405 CTCTGAGATCAACACCATCTTAGTAGTGAAGCTCTTATGCAATGGATTCGCACACAGAA 464  
115 nHisPhe---LysArgAlaArgLeuPheGluAspGluAspGluAspAlaGluValI 134  
465 TCAGTTTCAACAAACGTCGCTCGATTGTTTGAAGATCCTGAACCTCAAGATGCTAAGGTCT 524  
134 ePheProSerSerValTyrProSerProLeuProGluSerThrValProAlaAsnLysG 154  
525 TTTATCCATCGAAC-----CCTGAATCTACTGAACCGACGATGAATAAGG 566  
154 yTyr---AlaSerSerAlaIleGlnThrLeuPheThrGlyProValLysAlaGluL 173  
567 TTTATGCGGTTCTACAGCCATCCAAAGCTTTTTCAAA---GAATCTAAAGCTGAAGAA-- 621  
173 oThrProThrProLysIleProLysValArgGlyLysLysLysAsnAlaAspProG 193  
622 -----ACGCCAAGGTACTTAAGAGAGAGAGAGAGAGAGAGATCCTTAATCCCGA 674  
193 uGluIleAsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrG 213  
675 GGAAGTAACTCTTCACTCCCGGTGGAGATGACTCAGAGAACCGCTCAAGTTCTACGA 734  
213 uSerAlaSerAlaArgLysArgThrValThrAlaGluArgGluArgAlaIleAsnAl 233  
735 GAGTGTCTTCTGTAGAAAGAGAACTGTAACTGAGAGAGAGAGAGAGAGAGCCGTCAATGC 794  
233 aAlaLysThrPheGluProThrAsnProPheArgValValLeuArgProSerTyrI 253  
795 AGCCAAACATTCGACCAACCAATCTTACTTACAGTTGTTCTGCGACCATCATCT 854  
253 uTyrArgGlyCysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGly 273  
855 ATACAGAGGTTGCATCATGTACTTGGCATCTGGGTTTGTCTGAGAAATACCTAAGTGGAT 914  
273 eSerGlyPheIleLysValGlnLeuAlaGluLysGlnTyrProValArgCysLeuTyr 293  
915 ATCTGSGTTTCATCAAGCTCCAGCTCGGTGAGAAACAAATGGCCAGTGGGTGCTTACAA 974  
293 sAlaGlyArgAlaLysPheSerGlnGlyTyrTyrGluPheThrLeuGluAsnLeuG 313  
975 AGCAGGAGAGCTAAGTTTAGCCAAAGATGGTATGAGTTCCACACTCGAGAACAAATAG 1034  
313 yGluGlyAspValCysValPheGluLeuLeuArgThrArgPheValLeuLysValTh 333  
1035 CGAAGGAGATGATGTGTGTGTAGCTTACTCAGAACTCGGATTTCTGTTCTCGAAGTCC 1094  
333 rAlaPheArgValAsnGluTyrVal 341  
1095 CGCTTTCTGTTCTAATGATGATGTG 1119  
RESULT 9  
AAC37617  
ID AAC37617 standard; DNA; 1344 BP.  
XX AAC37617;  
AC AAC37617;  
XX 17-OCT-2000 (first entry)  
DT  
XX







Db	304	CTTCTCTCTGGTTTGCTGAGAGTACCTTAAGTGGATCTCCGGGTTCATCAAGTCCAG	363
Qy	281	LeuAlaGluLysGlnTrpProValArgCysLeuTyriLysAlaClyArgAlaLysPheSer	300
Db	364	CTTCCGGAGAAACAATGGCCCTGTTCGATGTCTCTACAAGCCGGGAGAGCCAAATTCAGT	423
Qy	301	GlnGlyTTrpTyrrGluPheThrLeuGluAsnLeuGlyGluCysValPhe	320
Db	424	CARGATGTTACGAATTCCTCTAGAGAACCACTTAGGAGAGGAGACGTCTGTGTGTTT	483
Qy	321	GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValLanGluTyrr	340
Db	484	GAGTGTCTCAGAACCAAGAGATTCCTGTTTGAAGTGCAGCAGCTTCGAGTCAACGAGTAC	543
Qy	341	Val 341	
Db	544	GTC 546	
RESULT 11			
ABZ12468			
ID	ABZ12468	standard; DNA; 681 BP.	
XX	ABZ12468;		
AC			
XX			
DT	21-JAN-2003	(first entry)	
XX			
DE	Arabidopsis thaliana	stress regulated gene SEQ ID NO 273.	
XX			
KW	Arabidopsis thaliana;	plant; gene; stress; transgenic; ds.	
XX			
OS	Arabidopsis thaliana.		
OS			
XX	WO200216655-A2.		
PN			
XX			
PD	28-FEB-2002.		
XX			
PF	24-AUG-2001; 2001WO-US026685.		
XX			
PR	24-AUG-2000; 2000US-0227866P.		
PR	26-JAN-2001; 2001US-0264647P.		
PR	22-JUN-2001; 2001US-0300111P.		
XX			
PA	(SCRI ) SCRIPPS RES INST.		
PA	(SYGN ) SYNGENTA PARTICIPATIONS AG.		
XX			
PI	Harper JF, Kreps J, Wang X, Zhu T;		
XX			
DR	WPI; 2002-304127/34.		
XX			
PT	Identifying a stress condition to which a plant cell has been exposed and		
PT	producing plants with increased tolerance to these abiotic stresses.		
XX			
PS	Claim 144; SEQ ID NO 273; 577pp + Sequence Listing; English.		
XX			
CC	The invention relates to identifying a stress condition to which a plant		
CC	cell has been exposed, comprising: (a) contacting nucleic acid		
CC	representative of expressed polynucleotides in the plant cell with an		
CC	array or probes representative of the plant cell genome; and (b)		
CC	detecting a profile of expressed polynucleotides in the plant cell		
CC	characteristic of a stress response. The method is useful in the		
CC	production of transgenic plants, cells and seeds and in producing plants		
CC	with increased tolerance to abiotic stress. The present sequence is that		
CC	of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used		
CC	in methods of the invention. Note: The sequence data for this patent is		
CC	not represented in the printed specification but is based on sequence		
CC	information supplied to Derwent by the European Patent Office		
SO	Sequence 681 BP: 217 A; 139 C; 162 G; 163 T; 0 U; 0 Other;		

Percent Similarity:	87.18%	Conservative:	14
Best Local Similarity:	81.20%	Mismatches:	18
Query Match:	52.67%	Indels:	12
DB:	6	Gaps:	5

US-10-088-187A-11 (1-341) x ABZ12468 (1-681)

Qy	110	MetAspSerAlaHisAsnHisPhe---	LysAtgAlaArgLeuPheGluAspLeuGluAsp	128
Db	7	ATGGATTCCGCACAGAAATCAGTTT	CAACAAACGGTCGCGATTGTTGAAGATCTCTGAATC	66
Qy	129	GluAspAlaGluValIlePheProSerSerVal	tyrProSerProLeuProGluSerThr	148
Db	67	AAAGATGCTAAGGTCATTATCCATCGAAC	-----CCTGAATCTACT	108
Qy	149	ValProAlaAsnLysGlyTyr---	AlaSerSerAlaIleGlnThrLeuPheThrGlyPro	167
Db	109	GAACCAAGTGAATAAAGGTTATGGCGG	TTCTACAGCCATCCAAAGCTTTTTCAAA	165
Qy	168	ValLysAlaGluGluProThrProThrPro	LysIleProLysLysArgGlyArgLysLys	187
Db	166	TCTAAAGCTGAAGAA-----	ACGCCCAAGGTACTTAAAGAGAGAGGAAGAA	216
Qy	188	LysAsnAlaAspProGluGluIleAsnSer	SerAlaProArgAspAspAspProGluAsn	207
Db	217	AAGATCCTTAATCCGAGGAAGTAAACT	TTCCTCAACTCCCGGTGGAGATGACTCAGAGA	276
Qy	208	ArgSerLysPheTyrGluSerAlaAlaLys	ThrPheGluProThrAsnProPheArgValVal	247
Db	277	CGCTCAAGTTCACAGAGAGTGCTTCT	CTGCTAGAAAGAGAACTGTAACTGCAGAGGAA	336
Qy	228	GluArgAlaIleAsnAlaAlaLysThrPhe	GluProThrAsnProPheArgValVal	247
Db	337	GAGAGAGCCGTCATCGAGCCAAACAT	TCGAACCAACAAATCCTTACTTTAGAGTTGTT	396
Qy	248	LeuArgProSerTyrLeuTyrArgGlyCys	IleMetTyrLeuProSerGlyPheAlaGlu	267
Db	397	CTGGAGCATCATATCTATACAGAGGT	TGCATCATGTACTTGCCTCTGGGTTTGCTGAG	456
Qy	268	LysTyrLeuSerGlyIleSerGlyPheIle	LysValGlnLeuAlaGluLysGlnTrpPro	287
Db	457	AAATACCTAAGTGGGATATCTGGTTT	TCATCAAGCTCCAGCTCGGTGAGAAAACAAT	516
Qy	288	ValArgCysLeuTyrLysAlaGlyArgAla	LysPheSerGlnGlyTrpTyrGluPheThr	307
Db	517	GTGAGGTGCCTCTCAAGACAGGAGAG	CTAAGTTTAGCCAAGGATGTTATGTGTTGAG	576
Qy	308	LeuGluAsnAsnLeuGlyGluGlyAspVal	CysValPheGluLeuLeuArgThrArgAsp	327
Db	577	CTCGAACAATATATAGGCGAGGAGATG	TATGTGTGTTGAGTACTCAGAACTCGGAT	636
Qy	328	PheValLeuLysValThrAlaPheArg	ValAsnGluTyrVal	341
Db	637	TTCGTTCTCGAAGTCACCGCCTTT	CGTGTCAATGAGTATGTG	678

RESULT 12  
 ABX62122/c  
 ID ABX62122 standard; DNA; 426 BP.  
 XX  
 AC ABX62122;  
 XX  
 DT 25-FEB-2003 (first entry)  
 XX  
 DE Arabidopsis thaliana expressed sequence related polynucleotide #237.  
 XX  
 KW Transgenic plant; plant; genetically modified cell; environmental stress;  
 KW ribozyme creation; disease resistance; stress tolerance;  
 KW fungicide screening; insecticide screening; gene; ds.  
 XX  
 OS Unidentified.  
 XX  
 PN US2002040490-A1.  
 XX



SQ Sequence 1361 BP; 353 A; 317 C; 371 G; 320 T; 0 U; 0 Other;  
 Alignment Scores: Length: 1361  
 Pred. No.: 298.50  
 Score: 44.73%  
 Percent Similarity: 44.73%  
 Best Local Similarity: 29.71%  
 Query Match: 16.62%  
 Indels: 31  
 Gaps: 10

US-10-088-187A-11 (1-341) x ADC03457 (1-1361)  
 QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAla----- 37  
 Db 133 CGCATACCGGACAAAGTTTCGAGCAGTTCATCAGACAGATGCACATGCACATGCAGAGGGGTC 192  
 QY 38 ----ValAlaLeuThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAsp 56  
 Db 193 GACCTGAACCTGAAGACACCGAGCGGTGAGACATGGCATGGGTGGGTGTCAGCAGAGTTGCC 252  
 QY 57 AsnLysIleTrpPheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIle 76  
 Db 253 AATGACCTTTCTTGAAGTCAGATGGGAGGTTTCGCTAAGGCTCATGAAGTGCAGGAG 312  
 QY 77 GlyTyrLeuLeuIlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsn 96  
 Db 313 AATGACCTCTGCTCTTCACTGACATGCATGGCAATGCTCTTCGAGGTCTTAATCTTCGAC 372  
 QY 97 LeuSerHisSerGlu----IleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsn 115  
 Db 373 CCAGTGGGTGGAGAAATCTCTCTCTTCCTTCGCGGCGGAGATGCGCAACACTTTCAC 432  
 QY 116 HisPheLysArgAlaArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePhe 135  
 Db 433 AATATGGTGGTCAAGCGCGGTGAGCAATACCTCTGATTCGAT---GATACTAGGTTG 489  
 QY 136 ProSerSerValTyrProSerProLeuProGluSerThrValProAlaAsnLysGlyTyr 155  
 Db 490 CCGTCTCAGTTTGTGGATCCCTCAAAAGGCTCTTACTTCAAGAAATGCACTGCG--- 546  
 QY 156 AlaSerSerAlaIleGlnThrLeuPheThrGlyProValLysAlaGluGluProThrPro 175  
 Db 547 -----AAACTAAATCAAGCAAGAG 567  
 QY 176 ThrProLysIleProLysLysArgGlyArgLysLysAsnAlaAspProGluGluIle 195  
 Db 568 CTTCTGAAACCAACCAAGCAGCAGCTGTCTATGTTAAGCTTGAGGCAACTGAGGAACAG 627  
 QY 196 AsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrGluSerAla 215  
 Db 628 AAGAGT-----GATGATGACACATGCACCTGAGCCCGGTTACTTCTTCGAGG 675  
 QY 216 SerAlaArgLysArgThrValThrAlaGluGluArgLysAlaIleAsnAlaLys 235  
 Db 676 ACTGCAAGTCAGCTAAAT-----GAGGATGAAAACAGGAAATCATGGAGCGTGT 726  
 QY 236 ThrPheGluProThrAsnProPheArgValValLeuArgProSerTyrLeuTyrArg 255  
 Db 727 TCCATCCGACCGGGGAATCCCGCATTCGTGGTGGTCTCTGTATGATGCCCATCTCCACGC 786  
 QY 256 ---GlyCysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeu---SerGlyIle 273  
 Db 787 AAGAACACTTCTTGACCACTCCCGAGTAATTTGACGTGATCATCTTCAGAGCAAGCCG 846  
 QY 274 SerGlyPheIleLysValGlnLeuAla---GluLysGlnTrpProValArgCysLeuTyr 292  
 Db 847 CGTGAAGGCTCTGCTAGCTAGCTAGCAGAGAACAGCAAGTGGCATGTCAGGTGCTACTAT 906  
 QY 293 LysAlaGlyArgAlaLysPheSer---GlnGlyTyrTyrGluPheThrLeuGluAsnAsn 311  
 Db 907 TCCAGCAGGACAGATGCTTCACTGCCAACGTTGGTCAAGTTTGTGAGGAGCAATGGG 966

QY 312 LeuGlyGluGlyAspValCysValPheGluLeuLeuArg 324  
 Db 967 CTGCGGAGGGCGAGCTGCTGCTGCTTTGAGCTGATCAAA 1005

RESULT 14  
 ADC03415  
 ID ADC03415 standard; DNA; 1041 BP.  
 XX  
 AC ADC03415;  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Rice flowering time-related gene #28.  
 XX  
 KW flowering time; rice; flower architecture; plant biotechnology;  
 KW commercial plant farming; agriculture; flowering-related protein; gene;  
 KW ds.  
 OS Oryza sativa.  
 PN W02003000904-A2.  
 XX  
 PD 03-JAN-2003.  
 XX  
 PF 24-JUN-2002; 2002WO-EP006968.  
 XX  
 PR 22-JUN-2001; 2001US-0300112P.  
 PR 26-SEP-2001; 2001US-0325277P.  
 PR 30-NOV-2001; 2001US-0334984P.  
 XX  
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.  
 XX  
 PI Sessions A, Briggs S, Cooper B, Goff SA, Moughamer T; Zhu T;  
 PI Glazebrook J, Katagiri F, Krepis J, Provart N, Ricke D, Zhu T;  
 XX  
 DR WPI; 2003-229340/22.  
 DR P-PSDB; ADC03416.  
 XX  
 PT New isolated nucleic acid molecule encoding a polypeptide modulating  
 PT flower architecture and flowering time, useful in the area of plant  
 PT biotechnology, and commercial plant farming and agriculture.  
 XX  
 PS Claim 4; SEQ ID NO 55; 323pp; English.  
 XX  
 CC The invention comprises the amino acid and coding sequences of proteins  
 CC involved in the control of flowering time in rice. The DNA and protein  
 CC sequences of the invention are useful for modulating flower architecture  
 CC and flowering time, the DNA and protein sequences are useful in the area  
 CC of plant biotechnology, commercial plant farming and agriculture. The  
 CC present DNA sequence encodes a rice flowering-related protein of the  
 CC invention.  
 XX  
 SQ Sequence 1041 BP; 303 A; 218 C; 261 G; 259 T; 0 U; 0 Other;  
 Alignment Scores: Length: 1041  
 Pred. No.: 1.83e-20  
 Score: 293.50  
 Percent Similarity: 45.43%  
 Best Local Similarity: 27.76%  
 Query Match: 16.34%  
 Indels: 25  
 Gaps: 11

US-10-088-187A-11 (1-341) x ADC03415 (1-1041)  
 QY 20 LeuArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAla 39  
 Db 58 ATTAGCATACCGGAGAAAGTTTCGCGACAGATTCAGTGTGTCAGATCCCAAGGGTTCAAC 117  
 QY 40 LeuThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIle 59  
 Db 118 CTGAAGACCCAGTGGCGGCAACATGGGTGTCAGCGTCGAAAGGTTGCTGATGAGCTG 177  
 QY 60 TrpPheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeu 79

```
Db 178 ATCTCTCATGTCAGGATGGAGGATTTTGGCAAGGCTCATGAATTCGAGGAGATGACCTC 237
Qy 80 LeuilePhe-----ArgTyrGluGlyAsnSerAlaPheSerValTyrilePhe 95
Db 238 CTGTCTTTCACATGCAATGGCGTTCATATGGCAGGCTCTCTTTGATGATCTAATCTTT 297
Qy 96 AsnLeuSerHisSerGlu---TleAntyHisSerThrGlyLeuMetAspSerAlaHis 114
Db 298 GATGCAAGTGGCTGTGAGAAAGTCTCTCTCTTTATTGGTAAAGAACACAGTATAGT 357
Qy 115 AsnHisPheLysArgAlaLeuPheGluAspLeuGluAspGluAspAlaGluValle 134
Db 358 GGTGACACTCAGAAATGCTGAGCAATACCACTTTAAGTGATCTCAAGATACTAGCACA--- 414
Qy 135 PheProSerSer---ValTyrProSerProLeuProGluSerThrValProAlaAsnLys 153
Db 415 ----CCATCAATCTTCTAGTGTCTCTCAAGGCGCTCTACTTCAAGAAATGGAAT 471
Qy 154 GlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyProValLysAlaGluPro 173
Db 472 GCGACAGACCAAG-----ACTAATCCAAATAAA---GATGATCCA 507
Qy 174 ThrProThrProLysIleProLysLysArgGlyArgLysLysAsnAlaAspProGlu 193
Db 508 TTTTCAACAGGAAGAGCGCTGAAGATCCAAACAGCAGCGCTCTCATGTCAAAACATGAG 567
Qy 194 GluIleAsnSerSerAlaProArgAspAspProGlu-----AsnArgSerLys 210
Db 568 ATGATTGAGGAGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 627
Qy 211 PheTyrGluSerAlaSerAlaArgLysArgThrValThrAlaGluGluArgGluArgAla 230
Db 628 TACTACTACTCGAGGTTTGGC-----AATTACTTAACCTGGCGAGGAGTGAAGAG--- 678
Qy 231 IleAsnAlaLysThrPheGluProThrAsnProPheArgValValLeuArgPro 250
Db 679 ATATTCAAGTTGGTGTCTCACTCAACAGCAGCAATCTGTATTGTGACTGTTCTGCGAGCG 738
Qy 251 SerTyrLeuTyrArgGlyCysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeu 270
Db 739 CCACAAATTCATCGCAAGGCGCTCTGATCGTCCCGAGGATTTGACGCTGATCATCTT 798
Qy 271 SerGlyIleSerGlyPheIleLysValGlnLeuAlaGluLys-----GlnTrpProVal 288
Db 799 GACAGCAGATCACAGAAATCTGCTGATGAGACCAACAAAGAGAGAAATGGTATGTC 858
Qy 289 ArgCysLeuTyrLysAlaGlyArgAlaLysPheSer---GlnGlyTyrTrpGluPheThr 307
Db 859 AAATACTACCATGCAAGCACTACCGAGGCGTTCAATTGCAACGCTGGATCAAGTTTCATC 918
Qy 308 LeuGluAsnAsnLeuGlyGluGlyAspValCysValPheGluLeuLeuArg 324
Db 919 CCGGAGAACAGGCTGGAGAGGGTATACATCTGATCTTCGAGCTGATGAAA 969

RESULT 15
AAC35159
ID AAC35159 standard; DNA; 796 BP.
XX
AC AAC35159;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9203.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
```

```
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 200BP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
```

us-10-088-187a-11.rng

Mon Jan 3 13:13:30 2005

PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	08-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144332P.	PR	12-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144334P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159329P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159637P.
PR	21-JUL-1999;	99US-0144814P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0145086P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160741P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160814P.
PR	23-JUL-1999;	99US-0145218P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160980P.
PR	26-JUL-1999;	99US-0145276P.	PR	22-OCT-1999;	99US-0160981P.
PR	27-JUL-1999;	99US-0145913P.	PR	22-OCT-1999;	99US-0160989P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145919P.	PR	25-OCT-1999;	99US-0161405P.
PR	28-JUL-1999;	99US-0145951P.	PR	25-OCT-1999;	99US-0161406P.
PR	02-AUG-1999;	99US-0146386P.	PR	25-OCT-1999;	99US-0161359P.
PR	02-AUG-1999;	99US-0146388P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146389P.	PR	26-OCT-1999;	99US-0161361P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161920P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161992P.
PR	04-AUG-1999;	99US-0147302P.	PR	28-OCT-1999;	99US-0161993P.
PR	05-AUG-1999;	99US-0147192P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	08-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
QY	6 PheHisLysLeuIlePheSerSerThrIleGlnLysArgLeuArgValProAspLys	25	QY	6 PheHisLysLeuIlePheSerSerThrIleGlnLysArgLeuArgValProAspLys	25
DB	113 TTCTTCAAGTTGGTCTTACCTTCAACGATGAAGATAGATGATGAGGATTCGCCAAG	172	DB	113 TTCTTCAAGTTGGTCTTACCTTCAACGATGAAGATAGATGATGAGGATTCGCCAAG	172
QY	26 PheValSerLysPheLysAspGluLeuSerValAlaValAlaLeuThrValProAspGly	45	QY	26 PheValSerLysPheLysAspGluLeuSerValAlaValAlaLeuThrValProAspGly	45
DB	173 TTCGTAAGCTACAAGGCTCTAACTCTCAGAGGTTGTTACATTAGTGACTCTCTGCGGA	232	DB	173 TTCGTAAGCTACAAGGCTCTAACTCTCAGAGGTTGTTACATTAGTGACTCTCTGCGGA	232
QY	46 HisValTrpArgValGlyLeuArgLysAlaAspLysIleTrpPheGlnAspGlyTrp	65	QY	46 HisValTrpArgValGlyLeuArgLysAlaAspLysIleTrpPheGlnAspGlyTrp	65
DB	233 TATAAGCGTTCCATAAAGCTTAAACGGATCGGTGAAGAGATTGGTTCATGAGGTTGG	292	DB	233 TATAAGCGTTCCATAAAGCTTAAACGGATCGGTGAAGAGATTGGTTCATGAGGTTGG	292
QY	66 GlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeuIlePheArgTyrGlu	85	QY	66 GlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeuIlePheArgTyrGlu	85
DB	293 AGTGAGTTGCGAGGCTCATTCATCCATCGAGGAGGCCACTTCTCTTGTTCGAATACAAG	352	DB	293 AGTGAGTTGCGAGGCTCATTCATCCATCGAGGAGGCCACTTCTCTTGTTCGAATACAAG	352
QY	86 GlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSerGluIleAsnTyrHis	105	QY	86 GlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSerGluIleAsnTyrHis	105
DB	353 AAGAATTCGAGTTCCGTTGTAATAATCTTCAATGCTTCGCTTCGACGACCAACTATCCT	412	DB	353 AAGAATTCGAGTTCCGTTGTAATAATCTTCAATGCTTCGCTTCGACGACCAACTATCCT	412
QY	106 SerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArgLeuPheGluAsp	125	QY	106 SerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArgLeuPheGluAsp	125
DB	413 -----CTGGATGACGTTTCAT	427	DB	413 -----CTGGATGACGTTTCAT	427

```
QY 126 LeuGluAspGluAspAlaGluValIlePheProSerSerValTyrProSerProLeuPro 145
Db ::: ||| ||| ::::|||
428 ATCATCGACAGCGACGATGATATATC----- 454
QY 146 GluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThr 165
Db 454 ----- 454
QY 166 GlyProValLysAlaGluGluProThrProThrProLysIleProLysLysArgGlyArg 185
Db ||| ||| ||| ::|
455 -----GAAATTACTGGCAAAGAGTTTGTGGAACTCAGGGTACAAGG 496
QY 186 LysLysLys-AsnAlaAspProGluGluIleAsnSerSerAlaPro---ArgAspAspAs 204
Db ||| ||| ||| ::|
497 AAAAATAATCAATCTGTTAATGGTGTGTATACAGACACAAATCCAAAGAGAGACCAAGA 556
QY 204 pProGluAsnArgSerLysPheTyr-----GluSerAlaSerAlaAr 218
Db : ::| ::| ::| ::| ::| ::|
557 GACATCGAGTTTGACAAAGATCTACACGATGTTGATGTCAGAGTCTTAAAGAAAGAA 616
QY 218 gLysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThr 236
Db :|| ::| ::| ::| ::| ::|
617 GAAGAAGACAAGAGAGTCTCTCAGGGGATAACAGAGCTTTTAACAGACCGAGACC 671
```

Search completed: December 30, 2004, 06:32:39  
Job time : 581 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 30, 2004, 06:16:43 ; Search time 218 Seconds  
(without alignments)

1111.830 Million cell updates/sec

Title: US-10-088-187A-11

Perfect score: 1796

Sequence: 1 MRPFPFKLIFSSTIQEKL.....LLRTRDFVLKVTAFRVNEYV 341

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n model  
-Q=/cn2\_1/USFTO\_spool/US1008187/runat\_29122004\_115843\_147/app\_query.fasta\_1.519  
-DB=issued Patents NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US1008187 @CN 1.1 69 @runat\_29122004\_115843\_147 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:  
1: /cn2\_6/prodata/1/ina/5A\_COMB.seq:  
2: /cn2\_6/prodata/1/ina/5B\_COMB.seq:  
3: /cn2\_6/prodata/1/ina/6A\_COMB.seq:  
4: /cn2\_6/prodata/1/ina/6B\_COMB.seq:  
5: /cn2\_6/prodata/1/ina/PCUTUS\_COMB.seq:  
6: /cn2\_6/prodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	5.8	2269	3	US-09-351-457-3
2	104	5.8	2269	3	US-09-561-500-3
3	104	5.8	2269	3	US-09-561-108-3
4	104	5.8	2269	4	US-09-351-543-3
5	104	5.8	2269	4	US-09-561-526-3
6	104	5.8	2269	4	US-09-561-499-3
7	104	5.8	2269	4	US-09-998-831-3
8	104	5.8	2269	4	US-09-561-005-3
9	104	5.8	2269	4	US-08-819-386-3
10	104	5.8	2282	1	US-08-373-579-5
11	104	5.8	2282	1	US-08-418-595-5
12	104	5.8	2282	2	US-08-665-926-5

13	104	5.8	2282	3	US-09-162-437-5	Sequence 5, Appl1
14	104	5.8	2282	3	US-08-740-223A-5	Sequence 5, Appl1
15	104	5.8	2282	4	US-08-817-318-5	Sequence 5, Appl1
16	104	5.8	2282	4	US-09-709-188-5	Sequence 5, Appl1
17	104	5.8	2282	4	US-09-442-717-5	Sequence 5, Appl1
18	104	5.8	2282	4	US-09-689-020-5	Sequence 5, Appl1
19	97.5	5.4	3537	4	US-09-543-681A-3647	Sequence 3647, Ap
20	97.5	5.4	5751	4	US-09-023-655-1415	Sequence 1415, Ap
21	97.5	5.4	11907	3	US-08-061-376-4	Sequence 4, Appl1
22	97.5	5.4	14255	1	US-08-320-559-1	Sequence 1, Appl1
23	97.5	5.4	14255	1	US-08-327-392-1	Sequence 1, Appl1
24	97.5	5.4	14255	1	US-08-306-691B-55	Sequence 55, Appl1
25	97.5	5.4	14255	3	US-08-545-860D-1	Sequence 1, Appl1
26	97.5	5.4	14255	5	PCT-US94-04496-1	Sequence 1, Appl1
27	97	5.4	1200	4	US-09-107-532A-2744	Sequence 2744, Ap
28	96.5	5.4	4899	4	US-09-799-451-714	Sequence 714, App
29	95.5	5.3	3226	3	US-08-539-205A-5	Sequence 5, Appl1
30	95.5	5.3	3226	4	US-09-392-163A-5	Sequence 5, Appl1
31	94.5	5.3	1494	4	US-09-248-796A-1806	Sequence 1806, Ap
32	94.5	5.3	2082	4	US-09-248-796A-12787	Sequence 12787, A
c	33	94.5	6156	4	US-10-204-708-60	Sequence 60, Appl
34	93.5	5.2	825	4	US-09-489-039A-3190	Sequence 3190, Ap
35	92	5.1	2247	4	US-09-270-767-12793	Sequence 12793, A
c	36	92	71989	3	US-09-443-501A-2	Sequence 2, Appl1
37	91.5	5.1	3661	4	US-09-762-481B-1	Sequence 1, Appl1
38	91.5	5.1	7888	4	US-09-919-497-48	Sequence 48, Appl
c	39	91.5	8312	4	US-09-620-312D-1048	Sequence 1048, Ap
40	91	5.1	1420	1	US-08-080-255-3	Sequence 3, Appl1
41	91	5.1	1420	3	US-08-463-713-3	Sequence 3, Appl1
42	91	5.1	1420	5	PCT-US93-05857-3	Sequence 3, Appl1
43	91	5.1	3131	3	US-09-035-648-23	Sequence 23, Appl
44	91	5.1	3131	3	US-09-001-951-23	Sequence 23, Appl
45	91	5.1	3131	4	US-08-818-829-23	Sequence 23, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-351-457-3  
; Sequence 3, Application US/09351457  
; Patent No. 6312694  
; GENERAL INFORMATION:  
; APPLICANT: THORPE, PHILIP E.  
; APPLICANT: RAN, SOPHIA  
; TITLE OF INVENTION: CANCER TREATMENT METHODS USING THERAPEUTIC CONJUGATES  
; TITLE OF INVENTION: THAT BIND TO AMINOPHOSPHOLIPIDS  
; FILE REFERENCE: 4001.002300  
; CURRENT APPLICATION NUMBER: US/09/351,457  
; CURRENT FILING DATE: 1999-07-12  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2269  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-351-457-3

Alignment Scores:  
Pred. No.: 0.0263 Length: 2269  
Score: 104.00 Matches: 58  
Percent Similarity: 37.55% Conservativeness: 34  
Best Local Similarity: 23.67% Mismatches: 92  
Query Match: 5.79% Indels: 61  
DB: 3 Gaps: 10

US-10-088-187A-11 (1-341) x US-09-351-457-3 (1-2269)

QY 45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64  
Db 1304 GGAGCGGGTGGACAAATATTTCAGCGAGTGAGGATGCGCGTGTTCAGAGGACT 1363  
QY 65 TrpGln-----Glu 67  
|||

Db 1364 TGAAGAATATAAAGTGGATTGGTAACCCCTTCAGGAGAAATATTGCTGGGAATGAG 1423  
Qy 68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84  
1424 TTTGTTTCGCAACTGACTAATCAGCAACGCTATGTCTTAAATAACACCTTTAAAGACTGG 1483  
Qy 85 GluGlyAanSerAlaPheSerValTyr--IlePheAanLeuSerHisSerGluIleAan 103  
1484 GAAGGAATGAGCTTACTCATGTTATGAACATTTCTATCTCAAGTGAAGAACTCAAT 1543  
Qy 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAanHisPheLysArgAlaArg 121  
1544 TATAGATTACCTTAAAGACTTACAGGACAGCGCGAAATAAGCAGCATCAGCCAA 1603  
Qy 122 LeuPheGluAspLeuGluAspAlaGlu-----ValIlePheProSer 137  
1604 CCAGGAATGATTATTTAGCACAAAGATGAGACAAACGACAAATGTTTCAAAATGTTCA 1663  
Qy 138 SerValTyr-----ProSerProLeu-ProGluSe 147  
1664 CAATGCTNACAGGAGCTGGTGTGTTGATGATGTGCTCTTCCAACTTGAACGGAATG 1723  
Qy 147 rThrValProAlaAanLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167  
1724 TACTATCCACAGGAGGAGCAACAAATAAGTTCAACGGCATTAAATGGTACTACTGGAAA 1783  
Qy 167 oValIysAlaGluGluProThrPro-----LysIleProLysLy 182  
1784 GGCTCAGGCTATTCGCTCAAGGCCAACCAACCATGATGATCCGACGAGATT----- 1835  
Qy 182 sArgGlyArgLysLysLysAanAlaAspProGluGluIleAanSerSerAlaProArgAs 202  
1836 -----TCTAAACATCCCACTGAGGAACTGTCTCGAATATTTTCAAGA 1885  
Qy 202 pAspAspProGluAanArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222  
1886 CTTAAGCCCAAGTGCAC-----TGAAAGTCACGGCTGCGCACTGT 1924  
Qy 222 lThrAlaGluGluArgGluArgAlaIleAan-----AlaAlaLysThrPh 237  
1925 GTCTCTTCCACACAGAGGGCGTGTGCTCGGTGTGAGGGACCCACATGCTCCAGATT 1984  
Qy 237 eGluProThrAan 241  
1985 AGAGCCTGTAAAC 1997

## RESULT 2

US-09-561-500-3  
; Sequence 3, Application US/09561500  
; Patent No. 6342219  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF  
; FILE REFERENCE: 4001.002500  
; CURRENT APPLICATION NUMBER: US/09/561,500  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR FILING DATE: 1999-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2269  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-561-500-3

Alignment Scores:  
Pred. No.: 0.0263 Length: 2269  
Score: 104.00 Matches: 58  
Percent Similarity: 37.55% Conservative: 34  
Best Local Similarity: 23.67% Mismatches: 92  
Query Match: 5.79% Indels: 61

DB: 3 Gaps: 10  
US-10-088-187A-11 (1-341) x US-09-561-500-3 (1-2269)  
Qy 45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAanLysIleTrpPheGlnAspGly 64  
1304 GGAGCGGGTGGCAATATTATTTCAGCGACGTGAGGATGCGCGTGTGTTTTCAGAGACT 1363  
Qy 65 TrpGln-----Glu 67  
1364 TGAAGAATATAAAGTGGATTGGTAACCCCTTCAGGAGAAATATTGCTGGGAATGAG 1423  
Qy 68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84  
1424 TTTGTTTCGCAACTGACTAATCAGCAACGCTATGTCTTAAATAACACCTTTAAAGACTGG 1483  
Qy 85 GluGlyAanSerAlaPheSerValTyr--IlePheAanLeuSerHisSerGluIleAan 103  
1484 GAAGGAATGAGCTTACTCATGTTATGAACATTTCTATCTCAAGTGAAGAACTCAAT 1543  
Qy 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAanHisPheLysArgAlaArg 121  
1544 TATAGATTACCTTAAAGACTTACAGGACAGCGCGAAATAAGCAGCATCAGCCAA 1603  
Qy 122 LeuPheGluAspLeuGluAspAlaGlu-----ValIlePheProSer 137  
1604 CCAGGAATGATTATTTAGCACAAAGATGAGACAAACGACAAATGTTTCAAAATGTTCA 1663  
Qy 138 SerValTyr-----ProSerProLeu-ProGluSe 147  
1664 CAATGCTNACAGGAGCTGGTGTGTTGATGATGTGCTCTTCCAACTTGAACGGAATG 1723  
Qy 147 rThrValProAlaAanLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167  
1724 TACTATCCACAGGAGGAGCAACAAATAAGTTCAACGGCATTAAATGGTACTACTGGAAA 1783  
Qy 167 oValIysAlaGluGluProThrPro-----LysIleProLysLy 182  
1784 GGCTCAGGCTATTCGCTCAAGGCCAACCAACCATGATGATCCGACGAGATT----- 1835  
Qy 182 sArgGlyArgLysLysLysAanAlaAspProGluGluIleAanSerSerAlaProArgAs 202  
1836 -----TCTAAACATCCCACTGAGGAACTGTCTCGAATATTTTCAAGA 1885  
Qy 202 pAspAspProGluAanArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222  
1886 CTTAAGCCCAAGTGCAC-----TGAAAGTCACGGCTGCGCACTGT 1924  
Qy 222 lThrAlaGluGluArgGluArgAlaIleAan-----AlaAlaLysThrPh 237  
1925 GTCTCTTCCACACAGAGGGCGTGTGCTCGGTGTGAGGGACCCACATGCTCCAGATT 1984  
Qy 237 eGluProThrAan 241  
1985 AGAGCCTGTAAAC 1997

## RESULT 3

US-09-561-108-3  
; Sequence 3, Application US/09561108  
; Patent No. 6342221  
; GENERAL INFORMATION:  
; APPLICANT: Philip E. Thorpe  
; APPLICANT: Rolf A. Brekken  
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF  
; FILE REFERENCE: 4001.002584  
; CURRENT APPLICATION NUMBER: US/09/561,108  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR FILING DATE: 1999-04-28  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2269

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-561-108-3

Alignment Scores:
Pred. No.: 0.0263 Length: 2269
Score: 104.00 Matches: 58
Percent Similarity: 37.55% Conservative: 34
Best Local Similarity: 23.67% Mismatches: 92
Query Match: 5.79% Indels: 61
DB: 10 Gaps: 10

US-10-088-187A-11 (1-341) x US-09-561-108-3 (1-2269)
QY 45 GlyHisValTrrArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64
Db 1304 GGAGCGCGGTGGACAATTATTTCAGCGACGTGAGGATGGCAGCGTGTGATTTCAGAGGACT 1363
QY 65 TrpGln-----Glu 67
Db 1364 TGGAAAGAAATATAAAGTGGGATTTGGTAAACCCCTTCAGGAGAATATTGGCTGGGAAATGAG 1423
QY 68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84
Db 1424 TTTGTTTCGCACTGACTAATCAGCAACGCTATGTCTTAAATATACACCTTAAAGACTGG 1483
QY 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
Db 1484 GAAGGGAATGAGCTTACTCTATGTATGAACATTTCTCTCTCAAGTGAAGAACTCAAT 1543
QY 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
Db 1544 TATAGGATTCACCTTAAAGGACTTACAGGACAGCGGCAAAATAGCAGCATCAGCCAA 1603
QY 122 LeuPheGluAspLeuGluAspAlaGlu-----ValIlePheProSer 137
Db 1604 CCAGGAAATGATTTTAGCAAAAGGATGGAGACAACGACAAATGATTTTGCATTTCA 1663
QY 138 SerValTyr-----ProSerProLeu-ProGluSe 147
Db 1664 CAATGCTAACAGGAGGCTGGTGTGATGATGCTGTGCTTCCAACTTGAACGGAATG 1723
QY 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
Db 1724 TACTATCCACAGGCGAGACACAAATAGTTCAACGGCATTAATGGTACTACTGGAAA 1783
QY 167 oValLysAlaGluGluProThrPro-----LysIleProLysLy 182
Db 1784 GGCTCAGGCTATTTCGCTCAAGGCCACACCATGATGATCCGACGACAGATT----- 1835
QY 182 sArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
Db 1836 -----TCTAAACATCCCGACTCCAGTCCAGGAACTGTCTCGAACTATTTTCAAAGA 1885
QY 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222
Db 1886 CTTAAGCCCGAGTCAC-----TGAAGTCAACGCTCGGCACTGT 1924
QY 222 lThrAlaGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237
Db 1925 GTCCTCTCCACACAGAGGGGCTGTGCTCGTGTGCTGACGGGACCCACATGCTCCAGATT 1984
QY 237 eGluProThrAsn 241
Db 1985 AGAGCCTGTAAC 1997

RESULT 4
US-09-351-543-3
; Sequence 3, Application US/09351543
; Patent No. 6406693
; GENERAL INFORMATION:
; APPLICANT: THORPE, PHILIP E.
; APPLICANT: RAN, SOPHIA
```

```

Db      1985 AGAGCCTGTAAC 1997
RESULT 5
US-09-561-526-3
; Sequence 3, Application US/09561526
; Patent No. 6416758
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002586
; CURRENT APPLICATION NUMBER: US/09/561,526
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-561-526-3
Alignment Scores:
Pred. No.: 0.0263      Length: 2269
Score: 104.00         Matches: 58
Percent Similarity: 37.55%      Conservative: 34
Best Local Similarity: 23.67%    Mismatches: 92
Query Match: 5.79%             Indels: 61
DB: 4                     Gaps: 10

US-10-088-187A-11 (1-341) x US-09-561-526-3 (1-2269)
Qy      45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64
Db      1304 GGAGCGGGTGGACAATTATTTCAGCGACGTGAGGATGGCAGCGTGTGATTTTCAGAGGACT 1363
Qy      65 TrpGln-----Glu 67
Db      1364 TGGAAAGAATATAAAGTGGGATTTGGTAACCTTCAGGAGAATATTGGCTGGGAATGAG 1423
Qy      68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84
Db      1424 TTTGTTCCCACTGACTTAATCAGCAACGCTATGCTTAAATATACACTTAAAGACTGG 1483
Qy      85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
Db      1484 GAAGGAATGAGGCTTACTCATTTGTATGAACATTTCTCTCAAGTGAAGAACTCAAT 1543
Qy      104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
Db      1544 TATAGATTTCACCTTAAAGGACTTACAGGACGCGGCAAAATAAGCAGCATCGACCAA 1603
Qy      122 LeuPheGluAspLeuGluAspAlaGlu-----ValIlePheProSer 137
Db      1604 CAAATGCTAACAGGAGGCTGGTGGTTTGATGATGCTGGTCTTCCAACTTGAACGGAATG 1723
Qy      147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
Db      1724 TACTATCCACAGGACGACACAAATAAGTTCAACGGCATTAATAGTGTACTACTGGAAA 1783
Qy      167 oValLysAlaGluGluProThrProThrPro-----LysIleProLysIy 182
Db      1784 GGCTCAGGCTATTTCGCTCAAGGCCACCAACATGATGATCGACACGAGATT----- 1835
Qy      182 sArgGlyA-rGlyLysAlaAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
Db      1836 -----TCTAAACATCCCAAGTCCACCTGAGGAAGTGTCTCGAACTATTTCGAAAGA 1885

```

```

Qy      202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222
Db      1886 CTTAAGCCAGTGCAC-----TGAAGTCACGGCTGGCGACTGT 1924
Qy      222 lThrAlaGluGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237
Db      1925 GTCCTCTCCACACACAGAGGCGGTGCTCGGTGCTGACGGGACCCACATGCTCCAGATT 1984
Qy      237 eGluProThrAsn 241
Db      1985 AGAGCCTGTAAC 1997
RESULT 6
US-09-561-499-3
; Sequence 3, Application US/09561499
; Patent No. 6524583
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002582
; CURRENT APPLICATION NUMBER: US/09/561,499
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-561-499-3
Alignment Scores:
Pred. No.: 0.0263      Length: 2269
Score: 104.00         Matches: 58
Percent Similarity: 37.55%      Conservative: 34
Best Local Similarity: 23.67%    Mismatches: 92
Query Match: 5.79%             Indels: 61
DB: 4                     Gaps: 10

US-10-088-187A-11 (1-341) x US-09-561-499-3 (1-2269)
Qy      45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64
Db      1304 GGAGCGGGTGGACAATTATTTCAGCGACGTGAGGATGGCAGCGTGTGATTTTCAGAGGACT 1363
Qy      65 TrpGln-----Glu 67
Db      1364 TGGAAAGAATATAAAGTGGGATTTGGTAACCTTCAGGAGAATATTGGCTGGGAATGAG 1423
Qy      68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84
Db      1424 TTTGTTCCCACTGACTTAATCAGCAACGCTATGCTTAAATATACACTTAAAGACTGG 1483
Qy      85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
Db      1484 GAAGGAATGAGGCTTACTCATTTGTATGAACATTTCTCTCAAGTGAAGAACTCAAT 1543
Qy      104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
Db      1544 TATAGATTTCACCTTAAAGGACTTACAGGACGCGGCAAAATAAGCAGCATCGACCAA 1603
Qy      122 LeuPheGluAspLeuGluAspAlaGlu-----ValIlePheProSer 137
Db      1604 CAAATGCTAACAGGAGGCTGGTGGTTTGATGATGCTGGTCTTCCAACTTGAACGGAATG 1723
Qy      138 SerValTyr-----ProSerProLeu-ProGluSe 147
Db      1664 CAAATGCTAACAGGAGGCTGGTGGTTTGATGATGCTGGTCTTCCAACTTGAACGGAATG 1723
Qy      147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
Db      1836 -----TCTAAACATCCCAAGTCCACCTGAGGAAGTGTCTCGAACTATTTCGAAAGA 1885

```



```

US-10-088-187A-11 (1-341) x US-09-819-386-3 (1-2269)
QY 45 GlyHisValTTPArgValGlyLeuArgLysAlaAspAsnLysIleTTPheGlnAspGly 64
Db 1304 GARGCGGGTGGCAATATTATTCAGCGACGTGAGGATGGCAGCTTATTTTCAGAGGACT 1363
QY 65 TTPGln-----Glu 67
Db 1364 TGAAGAATATAAAGTGGATTGGTAACCTTCAGGAGATAATTGGCTGGGAATAG 1423
QY 68 PheValAspArgTyrSerIleArgLysGlyTyrLeuLeu-----IlePheArgTyr 84
Db 1424 TTTGTTTCGCAACTGACTAATCAGCAACGCTATGCTTAAATAACACCTTAAAGACTGG 1483
QY 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
Db 1484 GAAGGGAATGAGGCTTACTCATCTGATGATGAACTTCTCTCAAGTGAAGAACTCAAT 1543
QY 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
Db 1544 TATAGGATTCACCTTAAAGGACTTCAGCGGACAGCGCGCAAAATAAGCAGCATCAGCCAA 1603
QY 122 LeuPheGluAspLeuGluAspAlaGlu-----ValIlePheProSer 137
Db 1604 CCAGGAATGATTTTAGCACAAAGGATGGAGACACGCAAAATGATTTGCAATGTTCA 1663
QY 138 SerValTyr-----ProSerProLeu-ProGluSe 147
Db 1664 CAATGCTAACAGGAGGCTGGTGGTTGATGATGCTGGTCTTCCAACTTGAACGGAATG 1723
QY 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
Db 1724 TACTATCCACAGAGGCGAGACACAAATAAGTTCAACGGCATTTAAATGGTACTACTGGAAA 1783
QY 167 oValLysAlaGluProThrPro-----LysIleProLysLy 182
Db 1784 GGCTCAGGCTATTCGCTCAAGGCCCAACCATGATGATCCGACGACAGATT----- 1835
QY 182 sArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
Db 1836 -----TCTAAACATCCAGTCCACCTCAGGAACTGTCTCGAACTATTTTCAAAGA 1885
QY 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222
Db 1886 CTTAAGCCCGAGTGCAC-----TGAAAGTCACGCGCTGGCAGCTGT 1924
QY 237 eGluProThrAsn 241
Db 1985 AGAGCCTGTAAC 1997

RESULT 10
US-08-373-579-5
; Sequence 5, Application US/08373579
; Patent No. 5650490
; GENERAL INFORMATION:
; APPLICANT: Davis, et al.
; TITLE OF INVENTION: TIE-2 LIGAND, METHOD OF MAKING AND USES
; NUMBER OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
;

US-09-819-386-3
; Sequence 3, Application US/09819386
; Patent No. 6783760
; GENERAL INFORMATION:
; APPLICANT: THORPE, PHILIP E.
; APPLICANT: RAN, SOPHIA
; TITLE OF INVENTION: CANCER TREATMENT METHODS USING THERAPEUTIC CONJUGATES
; TITLE OF INVENTION: THAT BIND TO AMINOPHOSPHOLIPIDS
; FILE REFERENCE: 4001.002300
; CURRENT APPLICATION NUMBER: US/09/819,386
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US/09/351,457
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-819-386-3

Alignment Scores:
Pred. No.: 0.0263 Length: 2269
Score: 104.00 Matches: 58
Percent Similarity: 37.55% Conservative: 34
Best Local Similarity: 23.67% Mismatches: 92
Query Match: 5.79% Indels: 61
DB: 4 Gaps: 10

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/373,579
; FILING DATE: 17-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/353,503
; FILING DATE: 09-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/348,492
; FILING DATE: 02-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/330,261
; FILING DATE: 27-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/319,932
; FILING DATE: 07-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Cobert, Robert J.
; REGISTRATION NUMBER: 36,108
; REFERENCE/DOCKET NUMBER: REG 330-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 345-7400
; TELEFAX: (914) 345-7721
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 357..1847
;
US-08-373-579-5
Alignment Scores:
Pred. No.: 0.0266 Length: 2282
Score: 104.00 Matches: 58
Percent Similarity: 37.55% Conservative: 34
Best Local Similarity: 23.67% Mismatches: 92
Query Match: 5.79% Indels: 61
DB: 1 Gaps: 10

US-10-088-187A-11 (1-341) x US-08-373-579-5 (1-2282)
Qy 45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64
Db 1311 GGAGCGGGTGGACAATTATTTCAGCGACGTGAGGATGGCGGTGTTGATTTTCAGAGGACT 1370
Qy 65 TrpGln-----Glu 67
Db 1371 TGGAAAGAAATATAAGTGGGATTGGTAACCTTCAGGAGAAATATTGCTGGGAATGAG 1430
Qy 68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84
Db 1431 TTTGTTTCGCACTGACTAATCAGCAACGCTATGCTTAAATACACTTAAAGACTGG 1490
Qy 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
Db 1491 GAAGGGAATGAGGCTTACTCATTTGTATGAACATTTCTATCTCTCAAGTGAAGAACTCAAT 1550
Qy 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 121
Db 1551 TATAGATTACCTTAAGGACTTACAGGACAGCGCGGCAAAATAAGCAGCATCAGCAA 1610
Qy 122 LeuPheGluAspLeuGluAspAlaGlu-----ValIlePheProSer 137
Db 1611 CCAGGAATGATTTAGCACAAAGATGGAGACACGACAAATGTTTGCAAATGTCA 1670
Qy 138 SerValTyr-----ProSerProLeu-ProGluSe 147
Db 1671 CAAATGCTAACAGGAGGCTGGTGGTTTGATGATGCTGCTTCAACTTGAACCGAATG 1730
Qy 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
Db 1731 TACTATCCACAGAGCGCAGACAAACAATAGTTCAACGGCATTAATGTTACTACTCTGAAA 1790
Qy 167 oValLysAlaGluGluProThrPro-----LysIleProLysLy 182
Db 1791 GGCTCAGGCTATTGCTCAAGGCCCAACACCATGATGATCGGACGACAGATT----- 1842
Qy 182 sArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
Db 1843 -----TCTAAACATCCACGAGTCCACCTGAGGAACTGCTCGAACTATTTTCAAGA 1892
Qy 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222
Db 1893 CTTAAGCCAGTGCAC-----TGAAGTCACGGCTCGGCACTGT 1931
Qy 222 lThrAlaGluGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237
Db 1932 GTCTCTCTTCCACACAGAGGCGGTGCTGCTCGGTGCTGACGGAGCCACATGCTCCAGATT 1991
Qy 237 eGluProThrAsn 241
Db 1992 AGAGCCTGTAAC 2004

RESULT 11
US-08-418-595-5
; Sequence 5, Application US/08418595
; Patent No. 5814464
; GENERAL INFORMATION:
; APPLICANT: Davis, et al.
; TITLE OF INVENTION: TIE-2 LIGAND, METHOD OF MAKING AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/418,595
; FILING DATE: 06-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,579
; FILING DATE: 17-JAN-1995
; APPLICATION NUMBER: US 08/353,503
; FILING DATE: 09-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/348,492
; FILING DATE: 02-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/330,261
; FILING DATE: 27-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/319,932
; FILING DATE: 07-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Cobert, Robert J.
; REGISTRATION NUMBER: 36,108
; REFERENCE/DOCKET NUMBER: REG 330-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 345-7400
; TELEFAX: (914) 345-7721
; INFORMATION FOR SEQ ID NO: 5:
```

SEQUENCE CHARACTERISTICS:  
LENGTH: 2282 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 357..1847  
US-08-418-595-5

Alignment Scores:  
Pred. No.: 0.0266 Length: 2282  
Score: 104.00 Matches: 58  
Percent Similarity: 37.55% Conservative: 34  
Best Local Similarity: 23.67% Mismatches: 92  
Query Match: 5.79% Indels: 61  
DB: 1 Gaps: 10

US-10-088-187A-11 (1-341) x US-08-418-595-5 (1-2282)

QY 45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64  
DB 1311 GGAGCGGGTGGACAATTATTCAGCGAGCTGCGAGGCTTATTTTCAGAGGACT 1370

QY 65 TrpGln-----Glu 67  
DB 1371 TGGAAAGATATAAGTGGGATTGGTTACCTTCAGGAGAAATATTGGCTGGGAAATGAG 1430

QY 68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84  
DB 1431 TTGTGTTGCGCACTGACTAATCAGCAACGCTATGTCTTAAATACACCTTAAGAGCTGG 1490

QY 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103  
DB 1491 GAAGGGAATGAGGCTTACTCTATTGTAACATTTCTATCTCTCAAGTGAAGAACTCAAT 1550

QY 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121  
DB 1551 TATAGGATTCACCTTAAAGGACTTACAGGACAGCGCGGCAAAAATAGCAGCATCAGCCAA 1610

QY 122 LeuPheGluAspLeuGluAspGluAspAlaGlu-----ValIlePheProSer 137  
DB 1611 CCAGGAATGATTTTATGACACAGGATGGAGACACGACAAATGTTATTCGAAATGTTCA 1670

QY 138 SerValTyr-----ProSerProLeu-ProGluSe 147  
DB 1671 CAAATGCTAAACAGGAGGCTGGTGGTTGATGCAATGGTCTCTTCAACTTGAACGGAATG 1730

QY 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167  
DB 1731 TACTATCCACAGGACGAGACACAAATAGTTCAACGGCATTAAATGTGTAATACTGAAA 1790

QY 167 oValLysAlaGluGluProThrPro-----LysIleProLysLy 182  
DB 1791 GGCTCAGGCTATTGCTCTCAAGGCCACACCATGATGATCCGACAGCAAT-----1842

QY 182 sArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202  
DB 1843 -----TCTAAACATCCCACTCCAGTCCACCTGAGAACTGTCTCGAATCTATTTTCAAGA 1892

QY 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222  
DB 1893 CITTAAGCCAGTGAC-----TGAAAGTCACGGCTGGCGCACTGT 1931

QY 222 lThrAlaGluGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237  
DB 1932 GTCTCTTCCACACAGAGGGCGGTGTGCTCGGTGTGACGGGACCCACATGCTCCAGATT 1991

QY 237 eGluProThrAsn 241  
DB 1992 AGAGCTGTAAAC 2004

RESULT 12  
US-08-665-926-5  
Sequence 5, Application US/08665926  
Patent No. 5851797  
GENERAL INFORMATION:  
APPLICANT: Valenzuela et al.  
TITLE OF INVENTION: TIE LIGAND-3, METHODS OF MAKING AND USES  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
STREET: 777 Old Saw Mill River Road  
CITY: Tarrytown  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10591-6707  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/665,926  
FILING DATE: 19-JUN-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Robert J. Cobert  
REGISTRATION NUMBER: 36,108  
REFERENCE/DOCKET NUMBER: REG 330-H  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (914) 345-7400  
TELEFAX: (914) 345-2113  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2282 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 357..1847  
US-08-665-926-5

Alignment Scores:  
Pred. No.: 0.0266 Length: 2282  
Score: 104.00 Matches: 58  
Percent Similarity: 37.55% Conservative: 34  
Best Local Similarity: 23.67% Mismatches: 92  
Query Match: 5.79% Indels: 61  
DB: 1 Gaps: 10

US-10-088-187A-11 (1-341) x US-08-665-926-5 (1-2282)

QY 45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64  
DB 1311 GGAGCGGGTGGACAATTATTCAGCGAGCTGCGAGGCTTATTTTCAGAGGACT 1370

QY 65 TrpGln-----Glu 67  
DB 1371 TGGAAAGATATAAGTGGGATTGGTTACCTTCAGGAGAAATATTGGCTGGGAAATGAG 1430

QY 68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84  
DB 1431 TTGTGTTGCGCACTGACTAATCAGCAACGCTATGTCTTAAATACACCTTAAGAGCTGG 1490

QY 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103  
DB 1491 GAAGGGAATGAGGCTTACTCTATTGTAACATTTCTATCTCTCAAGTGAAGAACTCAAT 1550

QY 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121  
DB 1551 TATAGGATTCACCTTAAAGGACTTACAGGACAGCGCGGCAAAAATAGCAGCATCAGCCAA 1610

QY 122 LeuPheGluAspLeuGluAspGluAspAlaGlu-----ValIlePheProSer 137  
DB 1611 CCAGGAATGATTTTATGACACAGGATGGAGACACGACAAATGTTATTCGAAATGTTCA 1670

QY 138 SerValTyr-----ProSerProLeu-ProGluSe 147  
DB 1671 CAAATGCTAAACAGGAGGCTGGTGGTTGATGCAATGGTCTCTTCAACTTGAACGGAATG 1730

QY 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167  
DB 1731 TACTATCCACAGGACGAGACACAAATAGTTCAACGGCATTAAATGTGTAATACTGAAA 1790

QY 167 oValLysAlaGluGluProThrPro-----LysIleProLysLy 182  
DB 1791 GGCTCAGGCTATTGCTCTCAAGGCCACACCATGATGATCCGACAGCAAT-----1842

QY 182 sArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202  
DB 1843 -----TCTAAACATCCCACTCCAGTCCACCTGAGAACTGTCTCGAATCTATTTTCAAGA 1892

QY 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222  
DB 1893 CITTAAGCCAGTGAC-----TGAAAGTCACGGCTGGCGCACTGT 1931

QY 222 lThrAlaGluGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237  
DB 1932 GTCTCTTCCACACAGAGGGCGGTGTGCTCGGTGTGACGGGACCCACATGCTCCAGATT 1991

QY 237 eGluProThrAsn 241  
DB 1992 AGAGCTGTAAAC 2004

122 LeuPheGluAspLeuGluAspGluAspAlaGlu-----ValIlePheProSer 137  
1611 CCAGGAATGATTTTAGCACAAGGATGGAGACACAAATGATTTGCAAAATGTTCA 1670  
138 SerValTyr-----ProSerProLeu-ProGluSe 147  
1671 CAAATCTAACAGGAGCGTGGTGGTTGATGCAATGGTCTTCCAACTTGAACGGAATG 1730  
147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167  
1731 TACTATCCACAGAGGAGCAGACAAATAAGTTCACCGCATTAATGGTACTACTGGAAA 1790  
167 oValLysAlaGluGluProThrPro-----LysIleProLysly 182  
1791 GGCTCAGGCTATTCGCTCAAGGCCACACCATGATCGACCGACGATTT-----1842  
182 sArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202  
1843 -----TCTAAACATCCCACTGAGGAACTGTCTCGAACTATTTTCAAGA 1892  
202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222  
1893 CTTAAGCCAGTGCAC-----TGAAGTCAAGCTGGCGACTGT 1931  
222 lThrAlaGluGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237  
1932 GTCTCTCCACACAGAGGCGGTGTCTCGGTGCTGCGGAGCCACATGCTCCAGATT 1991  
237 eGluProThrAsn 241  
1992 AGAGCCTGTAAC 2004

RESULT 13  
US-09-162-437-5  
Sequence 5, Application US/09162437  
Patent No. 6186185  
GENERAL INFORMATION:  
APPLICANT: Davis, et al.  
TITLE OF INVENTION: TIE-2 LIGAND, METHOD OF MAKING AND USES  
TITLE OF INVENTION: THEREOF  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
STREET: 777 Old Saw Mill River Road  
CITY: Tarrytown  
STATE: New York  
COUNTRY: USA  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/162,437  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/418,595  
FILING DATE: 06-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/373,579  
FILING DATE: 17-JAN-1995  
APPLICATION NUMBER: US 08/353,503  
FILING DATE: 09-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/348,492  
FILING DATE: 02-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/330,261  
FILING DATE: 27-OCT-1994  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/319,932  
FILING DATE: 07-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Cobert, Robert J.  
REGISTRATION NUMBER: 36,108  
REFERENCE/DOCKET NUMBER: REG 330-D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (914) 345-7400  
TELEFAX: (914) 345-7721  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2282 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 357..1847  
US-09-162-437-5

Alignment Scores:  
Pred. No.: 0.0266 Length: 2282  
Score: 104.00 Matches: 58  
Percent Similarity: 37.55% Conservative: 34  
Best Local Similarity: 23.67% Mismatches: 92  
Query Match: 5.79% Indels: 61  
DB: 3 Gaps: 10

US-10-088-187A-11 (1-341) x US-09-162-437-5 (1-2282)

QY 45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64  
DB 1311 GGAGCGCGGTGGACAATTATTTCAGCGACGTGAGGATGGCAGCGTTGATTTTCAGAGGACT 1370  
QY 65 TrpGln-----Glu 67  
DB 1371 TGGAAAGAATATAAAGTGGGATTTGGTAACCCCTTCAGGAGAAATATTGGCTGGGAAATGAG 1430  
QY 68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84  
DB 1431 TTGTTCGCACTGCTAATCAACGACGCTATGTCCTTAAATACACCTTAAAGACTGG 1490  
QY 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103  
DB 1491 GAAGGAGATGAGCTTACTCATTTGATGACATTTCTCTCTCAAGTGAAGACTCAAT 1550  
QY 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121  
DB 1551 TATAGGATTCACCTTAAGGACTTACAGGGACAGCGCGCAAAATAAGCAGCATCAGCCAA 1610  
QY 122 LeuPheGluAspLeuGluAspAlaGlu-----ValIlePheProSer 137  
DB 1611 CCAGGAATGATTTTAGCACAAGGATGGAGACACAAATGATTTGCAAAATGTTCA 1670  
QY 138 SerValTyr-----ProSerProLeu-ProGluSe 147  
DB 1671 CAAATGCTAACAGGAGCGTGGTGGTTGATGCAATGGTCTTCCAACTTGAACGGAATG 1730  
QY 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167  
DB 1731 TACTATCCACAGAGGAGCAGACAAATAAGTTCACCGCATTAATGGTACTACTGGAAA 1790  
QY 167 oValLysAlaGluGluProThrPro-----LysIleProLysly 182  
DB 1791 GGCTCAGGCTATTCGCTCAAGGCCACACCATGATCGACCGACGATTT-----1842  
QY 182 sArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202  
DB 1843 -----TCTAAACATCCCACTGAGGAACTGTCTCGAACTATTTTCAAGA 1892  
QY 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222

```
Db 1893 CTTAAGCCAGTGCAC-----TGAAGTCACGGCTGGCGACTGT 1931
Qy 222 lThrAlaGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237
Db 1932 GTCTCTTCCACACAGAGGGGGTGTCTCGTGTGTGACGGGACCCACATCTCCAGATT 1991
Qy 237 eGluProThrAsn 241
Db 1992 AGAGCCTGTAAAC 2004

RESULT 14
US-08-740-223A-5
; Sequence 5, Application US/08740223A
; Patent No. 628584
; GENERAL INFORMATION:
; APPLICANT: Davis, et al.
; TITLE OF INVENTION: Expressed Ligand - Vascular
; TITLE OF INVENTION: Intercellular Signalling Molecule
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS: 28
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/740,223A
; FILING DATE: 25-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 60/022/999
; FILING DATE: 02-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Covert, Robert J
; REGISTRATION NUMBER: 36,108
; REFERENCE/DOCKET NUMBER: REG 333
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-345-7400
; TELEFAX: 914-345-7721
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 357...1844
; OTHER INFORMATION:
; NAME/KEY: Human TIE-2 ligand 2
; LOCATION: 1...2282
; OTHER INFORMATION: from clone pbluescript KS
; OTHER INFORMATION: encoding human TIE 2 ligand 2
US-08-740-223A-5

Alignment Scores:
Pred. No.: 0.0266 Length: 2282
Score: 104.00 Matches: 58
Percent Similarity: 37.55% Conservative: 34
Best Local Similarity: 23.67% Mismatches: 92
Query Match: 5.79% Indels: 61
DB: 3 Gaps: 10

US-10-088-187A-11 (1-341) x US-08-740-223A-5 (1-2282)
Qy 45 GlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64
```

```
Db 1311 GGAGCGGGTGGACAATTATTTCAGCGACGTGAGGATGGCAGCGTGTGATTTTCAGAGGACT 1370
Qy 65 TrpGln-----Glu 67
Db 1371 TGGAAAGAATATAAAGTGGGATTGGTAAACCTTCAGGAGAAATATTTGGCTGGGAAATGAG 1430
Qy 68 PheValaspArgTyrSerIleArgIleGlyTyrLeu-----IlePheArgTyr 84
Db 1431 TTTGTTTCGCACTGATTAATCAGCAACGCTATGTCTTAAATAACACCTTAAAGACTGG 1490
Qy 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
Db 1491 GAAGGGAATCAGGCTTACTCATTTGTATGAACATTTCTCTCAAGTGAAGAACTCAAT 1550
Qy 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
Db 1551 TATAGGATTCACCTTAAAGGACTTACAGGACAGCCGGCAAAATAAGCAGCATCAGCCAA 1610
Qy 122 LeuPheGluAspLeuGluAspGluAspAlaGlu-----ValIlePheProSer 137
Db 1611 CCAGGAATGATTTTACCAACAAGGATGGAGACACGACAAATGATTTTCAAAATGTTC 1670
Qy 138 SerValTyr-----ProSerProLeu-ProGluSe 147
Db 1671 CAAATGCTAACAGAGGCTGGTGGTTTGTATGTCATGTGGTCTCTTCCAACCTTGAACGGG 1730
Qy 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
Db 1731 TACTATCCACAGGCGAGAACACAAATAAGTTCAACGGCATTAATAGTACTACTGAAA 1790
Qy 167 oValLysAlaGluProThrProThrPro-----LysIleProLysLy 182
Db 1791 GGCTCAGGCTATTTCGCTCAAGGCCAACCAATGATGATCCGACGAGATT----- 1842
Qy 182 sArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
Db 1843 -----TCTAAACATCCCACTCCAGTCCCACTGAGGAACCTGTCTCGAACTATTTTCAAGA 1892
Qy 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222
Db 1893 CTTAAGCCAGTGCAC-----TGAAGTCACGGCTGGCGACTGT 1931
Qy 222 lThrAlaGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237
Db 1932 GTCTCTTCCACACAGAGGGCGTGTCTCGTGTCTGACGGGACCCACATGCTCCAGATT 1991
Qy 237 eGluProThrAsn 241
Db 1992 AGAGCCTGTAAAC 2004

RESULT 15
US-08-817-318-5
; Sequence 5, Application US/08817318
; Patent No. 6433143
; GENERAL INFORMATION:
; APPLICANT: Davis, Samuel et al.
; TITLE OF INVENTION: TIE-2 LIGANDS, METHODS OF MAKING AND USES THEREOF
; FILE REFERENCE: REG 330-F-PCT-US
; CURRENT APPLICATION NUMBER: US/08/817,318
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 2282
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Unknown Organism
; NAME/KEY: CDS
; LOCATION: (357)...(1847)
US-08-817-318-5
```

## Alignment Scores:

Pred. No.: 0.0266 Length: 2282  
Score: 104.00 Matches: 58  
Percent Similarity: 37.55% Conservative: 34  
Best Local Similarity: 23.67% Mismatches: 92  
Query Match: 5.79% Indels: 61  
DB: 4 Gaps: 10

US-10-088-187A-11 (1-341) x US-08-817-318-5 (1-2282)

```
QY 45 GlyHisValTyrArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGly 64
   |||
   |||
   |||
Db 1311 GGAGCGCGGTGACAAATATTACGCGAGGTGGAGGCTTGATTTTCAGAGGACT 1370

QY 65 TrpGln-----Glu 67
   |||
   |||
   |||
Db 1371 TGGAAAGATATAAAGTGGGATTGGTAACCTTCAGGAGAATATTGGCTGGGAATGAG 1430

QY 68 PheValAspArgTyrSerIleArgIleGlyTyrLeuLeu-----IlePheArgTyr 84
   |||
   |||
   |||
Db 1431 TTTGTTTCGCAACTGACTAATCAGCAACGCTATGTCTTAAATAACACCTTAAAGACTGG 1490

QY 85 GluGlyAsnSerAlaPheSerValTyr---IlePheAsnLeuSerHisSerGluIleAsn 103
   |||
   |||
   |||
Db 1491 GAAGGGAATGAGCTTACTCATTTGATGAACATTTCTATCTCAAGTGAAGAACTCAAT 1550

QY 104 Tyr-----HisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121
   |||
   |||
   |||
Db 1551 TATAGNTTCACCTTAAAGGACTTACAGGGCAGCGCGCAAAATAAGCAGCATCAGCCAA 1610

QY 122 LeuPheGluAspLeuGluAspGluAspAlaGlu-----ValIlePheProSer 137
   |||
   |||
   |||
Db 1611 CCAGGAATGATTTTAGCACAAAGGATGGAGACACGACAAATGTATTGCAATGTTCA 1670

QY 138 SerValTyr-----ProSerProLeu-ProGluSe 147
   |||
   |||
   |||
Db 1671 CAATGCTAACAGGAGGTGGTGGTTTGATGATGGTCTTCCAACTTGAACGGAATG 1730

QY 147 rThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGlyPr 167
   |||
   |||
   |||
Db 1731 TACTATCCACAGGCGCAGACACAAATAAGTTCAACGGCATTAATGGTACTACTGGAAA 1790

QY 167 oValLysAlaGluGluProThrProThrPro-----LysIleProLysLy 182
   |||
   |||
   |||
Db 1791 GGCTCAGGCTATTGCTCAAGGCCACACCATCATGATCCGACGACGATT----- 1842

QY 182 sArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAs 202
   |||
   |||
   |||
Db 1843 -----TCTAAACATCCCACTCCACCTGAGGAACCTGCTCGAACTATTTCAAAAGA 1892

QY 202 pAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrVa 222
   |||
   |||
   |||
Db 1893 CTTAAGCCAGTGCAC-----TGAAAGTCACGGCTCGGCACTGT 1931

QY 222 lThrAlaGluArgGluArgAlaIleAsn-----AlaAlaLysThrPh 237
   |||
   |||
   |||
Db 1932 GTCTCTTCCACACAGAGGCGGTGCTCGGTGCTGACGGACCCACATGCTCCAGATT 1991

QY 237 eGluProThrAsn 241
   |||
   |||
   |||
Db 1992 AGAGCCTGTAAC 2004
```

Search completed: December 30, 2004, 09:53:17  
Job time : 227 secs

**THIS PAGE BLANK (USPIS)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 30, 2004, 06:32:43 ; Search time 1535 Seconds  
(without alignments)  
1254.759 Million cell updates/sec

Title: US-10-088-187A-11

Perfect score: 1796  
Sequence: 1 MPRPFHKLIFSSTIQKRL.....LLRTRDFVLKVTAFRVNEYV 341

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4176236 seqs, 2824127955 residues

Total number of hits satisfying chosen parameters: 8352472

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US1008187/runat\_29122004.115844\_192/app\_query.fasta\_1.519  
-DB=Published Applications NA -QWTF=fastcap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USRR=US1008187@cgn\_1.1.480@runat\_29122004.115844\_192  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCk=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_1/ptodata/1/pubpna/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq\*
- 19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq\*
- 20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1084	60.4	1471	17	US-10-767-795-1318	Sequence 1318, Ap
2	946	52.7	681	9	US-09-938-842A-273	Sequence 273, App
3	946	52.7	681	11	US-09-938-842A-273	Sequence 273, App
4	908.5	50.6	2141	16	US-10-424-599-118105	Sequence 118105, Ap
5	880	49.0	1111	17	US-10-767-795-1316	Sequence 1316, Ap
6	835.5	46.5	1760	16	US-10-424-599-97592	Sequence 97592, A
7	640	35.6	1062	17	US-10-767-795-1317	Sequence 1317, Ap
8	582.5	32.4	644	17	US-10-021-323-9015	Sequence 9015, Ap
9	512	28.5	426	9	US-09-770-423-237	Sequence 237, App
10	375.5	20.9	628	16	US-10-424-599-97591	Sequence 97591, A
11	330.5	18.4	650	17	US-10-767-795-1320	Sequence 1320, Ap
12	329	18.3	2207	18	US-10-739-930-3743	Sequence 3743, Ap
13	303	16.9	2338	18	US-10-425-115-89685	Sequence 89685, A
14	281	15.6	1477	17	US-10-437-963-89631	Sequence 89631, A
15	280	15.6	1455	16	US-10-425-114-32044	Sequence 32044, A
16	278	15.0	1209	17	US-10-437-963-54928	Sequence 54928, A
17	269	15.0	2253	17	US-10-437-963-75700	Sequence 75700, A
18	268	14.9	769	16	US-10-424-599-23571	Sequence 23571, A
19	265	14.8	602	17	US-10-021-323-7926	Sequence 7926, Ap
20	255.5	14.3	1699	18	US-10-425-115-107372	Sequence 107372, A
21	246	13.7	2214	17	US-10-437-963-101386	Sequence 101386, A
22	241.5	13.4	986	17	US-10-437-963-49771	Sequence 49771, A
23	229.5	12.8	2000	9	US-09-938-842A-2973	Sequence 2973, Ap
24	229.5	12.8	2000	11	US-09-938-842A-2973	Sequence 2973, Ap
25	214	11.9	921	17	US-10-437-963-27186	Sequence 27186, A
26	210.5	11.7	494	16	US-10-424-599-64127	Sequence 64127, A
27	206.5	11.5	1242	17	US-10-767-795-2666	Sequence 2666, Ap
28	205.5	11.4	1578	18	US-10-425-115-130216	Sequence 130216, A
29	197	11.0	1560	17	US-10-437-963-84667	Sequence 84667, A
30	195.5	10.9	4068	17	US-10-437-963-35166	Sequence 35166, A
31	193.5	10.8	593	16	US-10-424-599-36556	Sequence 36556, A
32	193	10.7	1611	17	US-10-437-963-60237	Sequence 60237, A
33	186	10.4	1611	17	US-10-437-963-93094	Sequence 93094, A
34	183	10.2	634	16	US-10-424-599-89606	Sequence 89606, A
35	178	9.9	1617	17	US-10-437-963-18071	Sequence 18071, A
36	174	9.7	1269	17	US-10-437-963-74762	Sequence 74762, A
37	171.5	9.5	1419	18	US-10-739-930-4151	Sequence 4151, Ap
38	170.5	9.5	1176	17	US-10-437-963-35940	Sequence 35940, A
39	167	9.3	1152	16	US-10-425-114-3941	Sequence 3941, Ap
40	166	9.2	946	17	US-10-437-963-54432	Sequence 54432, A
41	165.5	9.2	646	16	US-10-424-599-99156	Sequence 99156, A
42	162	9.0	996	16	US-10-424-599-73748	Sequence 73748, A
43	161.5	9.0	1059	9	US-09-938-842A-312	Sequence 312, App
44	161.5	9.0	1059	11	US-09-938-842A-312	Sequence 312, App
45	154.5	8.6	631	18	US-10-425-115-89684	Sequence 89684, A

ALIGNMENTS

RESULT 1  
US-10-767-795-1318  
; Sequence 1318, Application US/10767795  
; Publication NO. US20040181830A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Zhou, Yihua  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53534)B  
; CURRENT APPLICATION NUMBER: US/10767,795  
; CURRENT FILING DATE: 2004-01-30  
; NUMBER OF SEQ ID NOS: 117596  
; SEQ ID NO 1318  
; LENGTH: 1471  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE:  
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C1540\_1  
US-10-767-795-1318

Alignment Scores:		2.14e-127	Length:	1471
Pred. No.:		1084.00	Matches:	222
Score:		72.48%	Conservative:	43
Percent Similarity:		60.43%	Mismatches:	63
Best Local Similarity:		60.36%	Indels:	38
Query Match:		17	Gaps:	8
DB:				
US-10-088-187A-11 (1-341) x US-10-767-795-1318 (1-1471)				
QY	1	MetProArgProPhePheHisLysLeuLeuPheSerSerThrIleGlnGluLysArgLeu	20	
DB	202	ATGCCACGCCCTTTTTCATTAAGCTTATCTCTCCACTCTCTCCAGCAGCAAGAACTG	261	
QY	21	ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu	40	
DB	262	AGGATACCTGATAACTTTATTAAAGAAATTCAGGATGAGCTTCTGTGCTGCCGCTCTC	321	
QY	41	ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp	60	
DB	322	ACCGTTCCTGATGGTCATCTTTGGCGGTGAGGAATAAAGAAAGTTGACAAACAGCGCTGG	381	
QY	61	PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu	80	
DB	382	TTTAAGGAGGGTGGCAGGAATTTGTCGAACGTTACTATATCCGTTGGTGGCTACTTTTG	441	
QY	81	IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer	100	
DB	442	ATTTTCAGATATGAAGGAATTCGCTTTTCAGTGTAGCATATTTGATTTGTACAACCTCT	501	
QY	101	GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla	120	
DB	502	GAAATAAATCATCAACAAATGCTCTGTGTGTACTCAATCAAT---CTCGGAAGACAA	558	
QY	121	ArgLeuPheGluAspLeuGluAspGluAlaGluValIlePheProSerSerValTyr	140	
DB	559	TATCCATTTGAAGACTTGAAGTAT-----GAATGTCTCTCCAGCG-----	603	
QY	141	ProSerProLeuPro-----	145	
DB	604	-----CTGCCTAATTTGTTGGTGGTCCAAACTTAACTGTATTAACCTGGAGTGA	654	
QY	146	GluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeu-----	163	
DB	655	GAAAGTCAACCATCACGCACCTAAGGGTGTAAATCAACCTATTCGGGTATGGCTTTT	714	
QY	164	-----PheThrGlyProValLys---AlaGluGluPro	173	
DB	715	CATTACTGTATTGAACCTTAAAGATTCCTGCTAATGAATGAATGGAATCTTTCAGGT	774	
QY	174	ThrProThrProLysLysArgGlyArgLysLysAsnAlaAspProGlu	193	
DB	775	GCAGAGCTACCAAACTG---AAAAAACCCAGGAGGAAAGACGAGGTTTGAGCGCTAGT	831	
QY	194	GluIleAsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrGlu	213	
DB	832	GAAGAGGATCATCTCTTGGACATGAAGATGATCAATGCTGTAATGATTTATGAA	891	
QY	214	SerAlaSerAlaArgLysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAla	233	
DB	892	AGTGCTTCCCGCAGGAAGAAATTTGACTGCTGTAAGAAAGAGAGAGAGGATTAATGCA	951	
QY	234	AlaLysThrPheGluProThrAsnProPhePheArgValValLeuArgProSerTyrLeu	253	
DB	952	GCCAAAGCATTTGAGCCCTACTAACCTTTCTCGAGGGTCTGCTGGACCGCTTATCA	1011	
QY	254	TyrArgGlyCysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyIle	273	
DB	1012	TACAGGGATGTATTATGACTTACTACCATCTGCTTTGCTGAGCAACATCTAAGCGGGIT	1071	
QY	274	SerGlyPheIleLysValGlnLeuAlaGlu---LysGlnTrpProValArgCysLeuTyr	292	

Db 277 CCTCAAGTCTACGAGAGTCTCTCTGTAGAAAGAGAACTGTAACTGCAGAGGAAGA 336  
Qy 228 GluArgAlaIleAsnAlaAlaLysThrPheGluProThrAsnProPhePheArgValVal 247  
Db 337 GAGAGCGCGTCAATGCAGCCHAAACATTCGAACCAACAAATCTTACTTTAGAGTTGTT 396  
Qy 248 LeuArgProSerTyrLeuTyrArgGlyCysIleMetTyrLeuProSerGlyPheAlaGlu 267  
Db 397 CTGCGACCATCATATCTATACAGAGGTTCATCATCTGCCATCTGGGTTTCTGAG 456  
Qy 268 LysTyrLeuSerGlyIleSerGlyPheIleValGlnLeuAlaGluLysGlnTrpPro 287  
Db 457 AAATACCTTAAGTGGGATATCTCTGTTTCATCAAGCTCCAGCTCGGTGAGAAACAATGGCCA 516  
Qy 288 ValArgCysLeuTyrLysAlaGlyArgAlaLysPheSerGlnGlyTyrPheThr 307  
Db 517 GTGAGTGCCTCTACAAAGCAGGAGAGCTAAGTTAGCCAGGATGATGATGATTCACA 576  
Qy 308 LeuGluAsnAsnLeuGlyGluGlyAspValCysValPheGluLeuLeuArgThrArgAsp 327  
Db 577 CTCGAGAACATATAGGCGAAGGAGATGTATGTGTGTTGAGCTACTCAGAACTCGGGAT 636  
Qy 328 PheValLeuLysValThrAlaPheArgValAsnGluTyrVal 341  
Db 637 TTCGTTCTCGAAGTCACCGCCTTTCGTGTCAATGATGATGTG 678

## RESULT 3

US-09-938-842A-273  
; Sequence 273, Application US/09938842A  
; Publication NO. US2004009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE OF INVENTION: SAME, AND METHODS OF USE  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 273  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-273

Alignment Scores:  
Pred. No.: 2,57e-110 Length: 681  
Score: 946.00 Matches: 190  
Percent Similarity: 87.18% Conservative: 14  
Best Local Similarity: 81.20% Mismatches: 18  
Query Match: 52.67% Indels: 12  
DB: 11 Gaps: 5

US-10-088-187A-11 (1-341) x US-09-938-842A-273 (1-681)

Qy 110 MetAspSerAlaHisAsnHisPhe---LysArgAlaArgLeuPheGluAspLeuGluAsp 128  
Db 7 ATGGATTCGCGCAGACAGATCAGTTCACAAACGTCGATGTTTGAAGATCCTGAATC 66  
Qy 129 GluAspAlaGluValPhePheSerSerValTyrProSerProLeuProGluSerThr 148  
Db 67 AAGATGCTAGTCAATTTATTCATCGAAC-----CCTGAATCTACT 108  
Qy 149 ValProAlaAsnLysGlyTyr---AlaSerSerAlaIleGlnThrLeuPheThrGlyPro 167

Db 109 GAACAGTGAATAAAGGTATTATGGCGGTCTCATAGCCATCCAAAGCTTTTTCAAA---GAA 165  
Qy 168 VallysAlaGluGluProThrProThrProLysLysPheProLysLysArgGlyArgLysLys 187  
Db 166 TCTAAGCTGAAGAA-----ACGCCAAGGTACTTAAGAGAGAGGAAGGAAGAAG 216  
Qy 188 LysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAspAspProGluAsn 207  
Db 217 AAGAATCCTAATCCGAGGAAGTAACTCTTCAACTCCCGGTGGAGATGACTCTCAGAGAAC 276  
Qy 208 ArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrValThrAlaGluGluArg 227  
Db 277 CGCTCAAAAGTCTTACGAGAGTCTTCTCTCTAGAAAGAGAACTGTAACTGCAGAGGAAGA 336  
Qy 228 GluArgAlaIleAsnAlaAlaLysThrPheGluProThrAsnProPhePheArgValVal 247  
Db 337 GAGAGCGCGTCAATGCAGCCAAACATTCGAAACCAACAAATCTTACTTTAGAGTTGTT 396  
Qy 248 LeuArgProSerTyrLeuTyrArgGlyCysIleMetTyrLeuProSerGlyPheAlaGlu 267  
Db 397 CTGCGACCATCATATCTATACAGAGGTTCATCATCTGCCATCTGGGTTTCTGAG 456  
Qy 268 LysTyrLeuSerGlyIleSerGlyPheIleLysValGlnLeuAlaGluLysGlnTrpPro 287  
Db 457 AAATACCTTAAGTGGGATATCTGTTTCATCAAGCTCCAGCTCGGTGAGAAACAATGGCCA 516  
Qy 288 ValArgCysLeuTyrLysAlaGlyArgAlaLysPheSerGlnGlyTyrPheThr 307  
Db 517 GTGAGTGCCTCTACAAAGCAGGAGAGCTAAGTTAGCCAGGATGATGATGATTCACA 576  
Qy 308 LeuGluAsnAsnLeuGlyGluGlyAspValCysValPheGluLeuLeuArgThrArgAsp 327  
Db 577 CTCGAGAACATATAGGCGAAGGAGATGTATGTGTGTTGAGCTACTCAGAACTCGGGAT 636  
Qy 328 PheValLeuLysValThrAlaPheArgValAsnGluTyrVal 341  
Db 637 TTCGTTCTCGAAGTCACCGCCTTTCGTGTCAATGATGATGTG 678

## RESULT 4

US-10-424-599-118105  
; Sequence 118105, Application US/10424599  
; Publication NO. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 118105  
; LENGTH: 2141  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_77658C.1  
US-10-424-599-118105

Alignment Scores:  
Pred. No.: 9,47e-105 Length: 2141  
Score: 208.50 Matches: 198  
Percent Similarity: 61.31% Conservative: 46  
Best Local Similarity: 49.75% Mismatches: 95  
Query Match: 50.58% Indels: 59  
DB: 16 Gaps: 8

US-10-088-187A-11 (1-341) x US-10-424-599-118105 (1-2141)

Qy 1 MetProArgProPhePheHisLysLeuIlePheSerSerThrIleGln---GluLysArg 19





Mon Jan 3 13:13:31 2005

US-10-767-795-1317  
 ; Sequence 1317, Application US/10767795  
 ; Publication No. US20040181830A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Zhou, Yihua  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
 ; FILE REFERENCE: 38-21(53534)B  
 ; CURRENT APPLICATION NUMBER: US/10/767,795  
 ; CURRENT FILING DATE: 2004-01-30  
 ; NUMBER OF SEQ ID NOS: 117596  
 ; SEQ ID NO 1317  
 ; LENGTH: 1062  
 ; TYPE: DNA  
 ; ORGANISM: Gossypium hirsutum  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C1540\_2  
 ; US-10-767-795-1317

Alignment Scores:  
 Pred. No.: 5,65e-71 Length: 1062  
 Score: 640.00 Matches: 131  
 Percent Similarity: 82.49% Conservative: 15  
 Best Local Similarity: 74.01% Mismatches: 20  
 Query Match: 35.63% Indels: 11  
 DB: 17 Gaps: 3

US-10-088-187A-11 (1-341) x US-10-767-795-1317 (1-1062)

QY 165 ThrGlyProValLysAlaGluGluProThrProThrProLysIleProLysLysArgGly 184  
 DB 141 TCAGGTGCTATGAACCCAGAA-----CCAAAAGCGTGGG 176

QY 185 ArgLysLysAsnAlaAspProGluGluIleAsnSerAlaProArgAspAsp 204  
 DB 177 AGAAGCGGAG---TTTGTCTTAACGTCGAGGATTCATCTGTCGACGTGAAGATGAT 233

QY 205 ProGluAsnArgSerLysPheThrGluSerAlaSerAlaArgLysArgThrValThrAla 224  
 DB 234 GCTGACATGCGCTTAGTGTCTATGAAGTGTCTTCTGCCGAAAGAGAACCGTGACAGCT 293

QY 225 GluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThrAsnProPhePhe 244  
 DB 294 GAAGAAGAGAGAGAGCAATCAATGGCGCAAGCAATTTGAGCCTACTTAACCTTTCTGC 353

QY 245 ArgValValLeuArgProSerThrLeuThrArgGlyCysIleMetThrLeuProSerGly 264  
 DB 354 AGGTGCTCTTGGACCATCATATCTGTACAGGGGATGATATGATGATCTTACCGTCTGTC 413

QY 265 PheAlaGluLysThrLeuSerGlyIleSerGlyPhe-IleLysValGluLeuAlaGlu-- 283  
 DB 414 TTTGCTGAGAAGCATCTTAAGTGGGGTTCTCGGATTCATTAACCTTCAGCTTCCTGATGG 473

QY 284 -LysGlnTrpProValArgCysLeuThrLysAlaGlyArgAlaLysPheSerGlnGlyTr 303  
 DB 474 GAGACAGTGTCTGTACATGTCGTATTAAGCAGGACGAGCTTAAGTTCACTCAGTCAGGATG 533

QY 303 pThrGluPheThrLeuGluAsnLeuGlyGluGlyAspValCysValPheGluLeuLe 323  
 DB 534 GTACGAATTTACATTTGAGAAATAATTTGGGAGAGGGGATGTCGTGTCTTTCAGCTGCT 593

QY 323 uArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGlu 339  
 DB 594 CAGATCGAGGGAATTCGTCTCAAGATTACCGTATTTTCGTGTAGGGA 642

RESULT 8  
 US-10-021-323-9015  
 ; Sequence 9015, Application US/10021323  
 ; Publication No. US20040123340A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Deikman, Jill

APPLICANT: Peng, Paul C.C.  
 ; APPLICANT: Fincher, Karen L.  
 ; APPLICANT: Ziegler, Todd E.  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants  
 ; FILE REFERENCE: 38-21(52274)B  
 ; CURRENT APPLICATION NUMBER: US/10/021,323  
 ; CURRENT FILING DATE: 2001-12-12  
 ; PRIOR APPLICATION NUMBER: US 60/255, 619  
 ; PRIOR FILING DATE: 2000-12-14  
 ; NUMBER OF SEQ ID NOS: 17880  
 ; SEQ ID NO 9015  
 ; LENGTH: 644  
 ; TYPE: DNA  
 ; ORGANISM: Gossypium hirsutum  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(644)  
 ; OTHER INFORMATION: unsure at all n locations  
 ; OTHER INFORMATION: Clone ID: LIB3828-022-Q1-K6-E11  
 ; US-10-021-323-9015

Alignment Scores:  
 Pred. No.: 5,57e-64 Length: 644  
 Score: 582.50 Matches: 131  
 Percent Similarity: 60.96% Conservative: 22  
 Best Local Similarity: 52.19% Mismatches: 50  
 Query Match: 32.43% Indels: 48  
 DB: 17 Gaps: 6

US-10-088-187A-11 (1-341) x US-10-021-323-9015 (1-644)

QY 22 ValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeuThr 41  
 DB 1 ATCCCCATTAACCTTTGTTAGAAATTTGGGACGAACTTCTTGTGTCTGTCTCTCACT 60

QY 42 ValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPhe 61  
 DB 61 GTTCCTGATGTCATGTTTGGCGTGTAGGAATAAGGAAAGCCGACACCAAGGTTTGGTTT 120

QY 62 GlnAspGlyTrpGlnGluPheValAspArgTrpSerIleArgIleGlyTrpLeuLeuIle 81  
 DB 121 CACGAGGTTTGGCAGGAATTTGTAGACGGTACCATATCCGTGTGGCTACTTACTTGGTT 180

QY 82 PheArgTrpGluGlyAsnSerAlaPheSerValTrpIlePheAsnLeuSerHisSerGlu 101  
 DB 181 TTCAGATATGAATGGAATTTCTTGTTCAGT----- 210

QY 102 IleAsnThrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla-Ar 121  
 DB 211 -----AMACCGTA 219

QY 121 gLeuPheGluAspLeuGluAspGluAspAla-----GluValIlePhePr 136  
 DB 220 TCCATTGAAAGCACTTGAAGATGACGATGATGATCTCTCCGACACTGCGAATTTGTTTTC 279

QY 136 oSerSerValTrpProSerProLeu-----ProGluSerThrValProAlaAsnLys 153  
 DB 280 TGGCTCTAAAGTTAATAACTGCACTGAGTGGTGAATACTCAATTTTCGTGGCTTAA 339

QY 153 sGlyThrAlaSerSerAlaIleGlnThrLeuPheThrGlyProValLysAlaGluGluPr 173  
 DB 340 GGGCATGATGAATGCCAATCT-----TTTCGAGGTGCACTG----- 373

QY 173 oThrProThrProLysIleProLysArgGlyArgLysLysLysAsnAlaAspProGl 193  
 DB 374 -----CTACCAAA---CCAAACAGCCTGGGAGGAAAGGAGAGGTTTGTATCACGC 423

QY 193 uGluIleAsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTrpGlu 213  
 DB 424 TGAACCAAGATTCATCTGTTGGACACGATATGATGATGTGTATCAAACTTTGATATATGA 483

QY 213 uSerAlaSerAlaArgLysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAl 233

Db 484 AAGTGTTCAGCGAAGAGAGAACTGTGCAACGGAAGAGAGAGAGCAATTAACGC 543  
Qy 233 aAlaLysThrPheGluProThrAsnProPheArgValValLeuArgProSerTyrLe 253  
Db 544 AGCCAATCGTTGAGCCGATGACCTTCTGCAGAGTTGCTTACGACCATCATATCT 603  
Qy 253 uTyrArgGlyCysIleMetTyrLeuProSer 263  
Db 604 ATACCGGGGATGATTATGTACTTACCATCA 634

## RESULT 9

US-09-770-423-237/c

; Sequence 237, Application US/09770423

; Publication No. US2002040490A1

; GENERAL INFORMATION:

; APPLICANT: Goralach, Jörn

; APPLICANT: An, Yong-Qiang

; APPLICANT: Hamilton, Carol M.

; APPLICANT: Price, Jennifer L.

; APPLICANT: Raines, Tracy M.

; APPLICANT: Yu, Yang

; APPLICANT: Rameaka, Joshua G.

; APPLICANT: Page, Amy

; APPLICANT: Matthew, Abraham V.

; APPLICANT: Ledford, Brooke L.

; APPLICANT: Woessner, Jeffrey P.

; APPLICANT: Haas, William David

; APPLICANT: Garcia, Carlos A.

; APPLICANT: Krieker, Maja

; APPLICANT: Slader, Ted

; APPLICANT: Davis, Keith R.

; APPLICANT: Allen, Keith

; APPLICANT: Hoffman, Neil

; APPLICANT: Hurban, Patrick

; TITLE OF INVENTION: Expressed Sequences of Arabidopsis

; FILE REFERENCE: 2028 (PARA-017PRV)

; CURRENT APPLICATION NUMBER: US/09/770,423

; CURRENT FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: 60/178,512

; PRIOR FILING DATE: 2000-01-27

; NUMBER OF SEQ ID NOS: 999

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 237

; LENGTH: 426

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; NAME/KEY: misc.feature

; LOCATION: (1)..(426)

; OTHER INFORMATION: n = A,T,C or G

US-09-770-423-237

Alignment Scores:  
Pred. No.: 2,87e-55 Length: 426  
Score: 512.00 Matches: 96  
Percent Similarity: 97.96% Conservative: 0  
Best Local Similarity: 97.96% Mismatches: 2  
Query Match: 28.51% Indels: 0  
DB: 9 Gaps: 0

US-10-088-187A-11 (1-341) x US-09-770-423-237 (1-426)

Qy 244 PheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyrLeuProSer 263  
Db 425 TTCAGAGTGGTTCGCGACCATCCTATCTATACAGAGGTGTCATGTATCTCTCTCT 366  
Qy 264 GlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGlnLeuAlaGlu 283  
Db 365 GGGTTGCTGAGAGTACCTAGTGGGATCTCCGGGTTTCATCAAGTCCAGNATGGGAG 306  
Qy 284 LysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSerGlnGlyTrp 303

Db 305 AAACAATGCGCTGTTTCGATGTCTCTACAAAGCGGAGAGCCAAATTCAGTCAGGATGG 246  
Qy 304 TyrGluPheThrLeuGluAsnLeuGlyGluGlyAspValCysValPheGluLeuLeu 323  
Db 245 TAGCAATTCACCTCTAGAGAACCACTTAGGAGAAGGAGAGCTCTGTGTGTTGAGCTGCTC 186  
Qy 324 ArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyrVal 341  
Db 185 AGAACCCAGAGATTTCTGTTTGAAGGTGACAGCCTTTCGAGTCACAGGATACGTC 132

## RESULT 10

US-10-424-599-97591

; Sequence 97591, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 97591

; LENGTH: 628

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_59138C.1

US-10-424-599-97591

Alignment Scores:  
Pred. No.: 1.37e-37 Length: 628  
Score: 375.50 Matches: 74  
Percent Similarity: 81.48% Conservative: 14  
Best Local Similarity: 68.52% Mismatches: 19  
Query Match: 20.91% Indels: 1  
DB: 16 Gaps: 1

US-10-088-187A-11 (1-341) x US-10-424-599-97591 (1-628)

Qy 182 LysArgGlyArgLysLysAsnAlaAspProGluGluLeuAsnSerAlaProArg 201  
Db 303 AAAACCGCAAAAAGAGCGGAATCAGAACCTATGGCGAGGAACCTTCTGGTGAAT 362  
Qy 202 AspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThr 221  
Db 363 GAAGAGGAGGCGAGAAATCGCTATAGGTTTATGAAAGTCATCTGCAAGGAACGAAAC 422  
Qy 222 ValThrAlaGluGluArgGluArgAlaIleAsnAlaLysThrPheGluProThrAsn 241  
Db 423 GTGACGGCAGAGAAAGAGAAAGGCAATTAATGTCATCAAAACCAATTGAAACCGCAAT 482  
Qy 242 ProPheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyrLeu 261  
Db 483 CCTTCTCCGAGTTGCTCTGGACCCCTCTATTTGATAGGGATGATATATGATCTG 542  
Qy 262 ProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGlnLeu 281  
Db 543 CCTTCCACCTTTGCTGAAAGAAATTTGAATGGGGTTTCGGGATTCATCAAACTTCAGCTG 602  
Qy 282 AlaGluLysGlnTrpProVal 288  
Db 603 TCCAATGTTAGACAGTGTGCGTT 626

## RESULT 11

US-10-767-795-1320

; Sequence 1320, Application US/10767795

; Publication No. US2004018130A1

; GENERAL INFORMATION:

APPLICANT: Kovalic, David K.  
APPLICANT: Cao, Yongwei  
APPLICANT: Zhou, Yihua  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
FILE REFERENCE: 38-21(53534)B  
CURRENT APPLICATION NUMBER: US/10/767,795  
CURRENT FILING DATE: 2004-01-30  
NUMBER OF SEQ ID NOS: 117596  
SEQ ID NO 1320  
LENGTH: 650  
TYPE: DNA  
ORGANISM: Gossypium hirsutum  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3543-054-Q1-K6-E2  
US-10-767-795-1320

Alignment Scores: Length: 650  
Pred. No.: 7,988-32 Matches: 70  
Score: 330.50 Conservative: 21  
Percent Similarity: 54.82%  
Best Local Similarity: 42.17%  
Query Match: 18.40%  
Indels: 56  
Gaps: 3  
DB:

US-10-088-187A-11 (1-341) x US-10-767-795-1320 (1-650)

QY 1 MetProArgProPhePheHisLysLeuLeuPheSerThrThrLeuGlnGluLysArgLeu 20  
DB 201 ATGCCACGCCCTTTTTCATTAAGCTTATCTCTACCCCTCCAAAGAAAG----- 254

QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40  
DB 254 ----- 254

QY 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60  
DB 255 -----AAACT-GGNAATAAGAAAGTTGACACAAAGGTTGG 289

QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTrpSerIleArgIleGlyTrpLeuLeu 80  
DB 290 TTTTCAGGAAGTTGGCAGGAGTTTCTAGACGCTTACTATATTCGTGGCTACACTG 349

QY 81 IlePheArgTrpGluGlyAsnSerAlaPheSerValTrpIlePheAsnLeuSerHisSer 100  
DB 350 GTTTTCAGATACCAAGAAATCTGCTTCAGTGTTAGTATATTTAAATTTGTATACTCG 409

QY 101 GluIleAsnTrpHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120  
DB 410 GAATAAAGTATCAGACTAATGCTTGGTGGTACTCAATCAATCAGCGAAACAATAT 469

QY 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTrp 140  
DB 470 CCG---TTTGAACAACCTTGAAGATGATGATGC----- 499

QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTrpAlaSerSerAlaIle 160  
DB 500 -----ATCTCTCCAGCACTT 514

QY 161 GlnThrLeuPheThrGly 166  
DB 515 CAGAAATTTGTTGGTGGG 532

RESULT 12  
US-10-739-930-3743/c  
Sequence 3743, Application US/10739930  
Publication No. US20040216190A1  
GENERAL INFORMATION:  
APPLICANT: Kovalic, David K.  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT  
FILE REFERENCE: 38-21(53377)B  
CURRENT APPLICATION NUMBER: US/10/739,930

CURRENT FILING DATE: 2003-12-18  
NUMBER OF SEQ ID NOS: 11088  
SEQ ID NO 3743  
LENGTH: 2207  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(2207)  
OTHER INFORMATION: unsure at all n locations  
FEATURE:  
OTHER INFORMATION: Clone ID: GLYMA-23APR03-CLUSTER63260\_1  
US-10-739-930-3743

Alignment Scores: Length: 2207  
Pred. No.: 8,498-31 Matches: 94  
Score: 329.00 Conservative: 57  
Percent Similarity: 43.77%  
Best Local Similarity: 27.25%  
Query Match: 18.32%  
Indels: 64  
Gaps: 10  
DB:

US-10-088-187A-11 (1-341) x US-10-739-930-3743 (1-2207)

QY 6 PheHisLysLeuLeuPheSerThrThrIleGlnGluLysArgLeuValProAspLys 25  
DB 1757 TTTTTCAGATCATATACCGCTCAATGTTTCATGAGAAAGCTAATGATCCCAATAAG 1698

QY 26 PheValSerLysPheLysAspGluLeuSerValAlaValAlaLeuThrValProAspGly 45  
DB 1697 TTTGTGAAGAAATATGGAAACCGCTTACAAATACCTCTATTCTGAGACTCCAAATGGT 1638

QY 46 HisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGlyTrp 65  
DB 1637 GCTCAATGGAAATGATTTTGAAGAAACGTCGCTAAATATGTTTCAAAAGGATGG 1578

QY 66 GlnGluPheValAspArgTrpSerIleArgIleGlyTrpLeuLeuLeuPheArgTrpGlu 85  
DB 1577 AAAGAGTTTCAGAGTATCATCTCTAGCTCATGGCCATCTTTTGGTTTTCAGATGGAT 1518

QY 86 GlyAsnSerAlaPheSerValTrpIlePheAsnLeuSerHisSerGluIleAsnTrpHis 105  
DB 1517 GTAACTTCTCATTTTCAGGTACACATCTTTGATTTGAGTGCCTTAGAGTTGAGTAC- 1461

QY 106 SerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArgLeuPheGluAsp 125  
DB 1461 ----- 1461

QY 126 LeuGluAspGluAspAlaGluValIlePheProSerSerValTrpProSerProLeuPro 145  
DB 1460 -----CCTACCGAAATA-----ATAAAA 1443

QY 146 GluSerThrValProAlaAsnLysGlyTrpAlaSerSerAlaIleGlnThrLeuPheThr 165  
DB 1442 GGTAAACAGCTCCATAGAAAGAAATGAATCTCCAGGTGATGAACATTTGGAGTGT 1383

QY 166 GlyProValLysAlaGluGluProThrProThrProLysIleProLysArgGlyArg 185  
DB 1382 -----CACAGATCAGGTCAA 1368

QY 186 LysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAspAsp----- 203  
DB 1367 AAGAAAGAAATTAATCTCAGTTGAATTTCTTCAACATGTCAATGAGAGTCCGACAGTGT 1308

QY 204 ---AspProGluAsn-----ArgSerLysPheTrpGluSerAlaSerAla 217  
DB 1307 GTTAAAGTTGAGACACGATCATCTGCAAGCAGGCTCTGCATCACACTGCCACAAA 1248

QY 218 ArgLys---ArgThrValThrAlaGluGluArgGluArgAlaIleAsnAlaLysThr 236  
DB 1247 TGCAAAAGAAATTCGAAGGCTATGATTAACCAAGTTTACAGCTCTTGTAGAGCAAGTCT 1188

QY 237 PheGluProThrAsnProPhePheArgValValLeuLeuArgProSerTrpLeuTrpArgGly 256

```
Db 1187 TTCAATCTGTAAATCCATCTCTTTCGCGTCATCATCGAACCCACATCATGTTCTCTAC 1128
Qy 257 CysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSer-----GlyIleSer 274
Db 1127 GGTAGTCTGAATCTACCAATGAAGTTCTCTAGGAGTCACCTTCATTTGCACAAGAGCGG 1068
Qy 275 GlyPheIleLysValGln---LeuAlaGluLysGlnTyrProValArgCys---LeuTyr 292
Db 1067 AGACTTATCAGCCTTCAGGCTTGAAGTGGAGAAATTTGGCTCGCAAAAGTATCAGATCCAT 1008
Qy 293 LysAlaGlyArgAla---LysPheSerGlnGlyTyrTyrGluPheThrLeuGluAsn 311
Db 1007 AAACAAAGACTGCTATAGATTAAACTCTCATGGAACGCATTTGTCAAGGACAATAAC 948
Qy 312 LeuGlyGluGlyAspValCysValPheGluLeuLeuLeuLeuLeuLeuLeuLeuLys 331
Db 947 TTGAAAGTTGGTGGATGTTGATCTTTGAACTCTTGAACCTCGTTCAATGGAACCTTAACCTTCCTG 888
Qy 332 ValThrAlaPheArg 336
Db 887 GTTCACATCTTCAGA 873
```

## RESULT 13

```
US-10-425-115-89685
; Sequence 89685, Application US/10425115
; Publication No. US20040124272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 89685
; LENGTH: 2338
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MR74577_181795C.1
US-10-425-115-89685
```

## Alignment Scores:

```
Pred. No.: 1,938-27 Length: 2338
Score: 303.00 Matches: 109
Percent Similarity: 40.64% Conservative: 56
Best Local Similarity: 26.85% Mismatches: 155
Query Match: 16.87% Indels: 86
DB: 18 Gaps: 17
```

US-10-088-187A-11 (1-341) x US-10-425-115-89685 (1-2338)

```
Qy 4 ProPhePheHisLysLeuIlePheSerThrIleGlnGluLysArgLeuArgValPro 23
Db 761 CCGCAGTCTTCAAGGTCTTCTTCCCGAGCAGTCCACGAG--CGGCTGAGATCCCA 817
Qy 24 AspLysPheValSerLysPheLysAspGluLeuSerVal---AlaValAlaLeuThrVal 42
Db 818 ACCATGTTCAACAGCACCTCAAGGAGCAGCAGCAACCGGAGCAGTTTCCCTGAGAGGC 877
Qy 43 ProAspGlyHisValTyrArgValGlyLeu-----ArgLysAlaAspAsnLysIleTyr 60
Db 878 CCGAGCGGCAACAGATGGCAGCAGCGCTCGTTCGAGTCCGAGTCGAGAGCCCGCGTGG 937
Qy 61 ---PheGlnAspGlyTyrGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeu 79
Db 938 TCGTTCGACAGGCTGGAGAGATTTGTACGGACCACTCCCTCAGGCTAGGCCACTTC 997
Qy 80 LeuIlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHis 99
```

```
Db 998 CTGGTCTTTCACGCGGAGCGGCCGCCCCAGTTCTCCGTGGCGGTGTCTCTCG----- 1048
Qy 100 SerGluIleAsnTyrHisSerThrGlyLeuMetAspSerAla----- 113
Db 1049 -----TCGTCAAGCGTTCATCGACCCCGGAGCCTCGACCGGAGGCT 1090
Qy 114 ---HisAsnHisPheLysArgAlaArgLeuPheGlu----- 124
Db 1091 ACCGCGAAGCGGCGCGCGCTCAAGCTCGAAGAGGCGAAGGCGTCCGCGTCCGAGGG 1150
Qy 125 AspLeuGlu-----AspGluAspAlaGluVal-----IlePheProSerSerValTyr 140
Db 1151 GAGCTGGAGCGAGCGCGCGACACCACTCCGAGGTGTCTCTCTCCCGCGGAAGAGGCG 1210
Qy 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160
Db 1211 GATGAGGAGCGACAGGAGGAGGAGACAGGAGCGAGTGTGGCGCTGCGCGCGCTCT 1270
Qy 161 GlnThrLeuPhe-----ThrGly----- 166
Db 1271 GAGATGTCTACTTGTCTTTAGAGAGAGGCGCGCTGAGTGACAGGGAAGAGAGCGAGCG 1330
Qy 167 -----ProValLysAlaGluGluProThrProThrProLysIleProLys 181
Db 1331 ACGACGAGTGATCTCCCGGCGAGATGCCAGTGCGCCAAGAAGACACTCTGCCCTCGCAAG 1390
Qy 182 LysArgGlyArgLysLysAsnAla----- 190
Db 1391 AAAGCTGTTAAAGAGGCGCTCAAGCTGCCACTTCAAAGGATGTGTGATGATGTTCTAC 1450
Qy 191 -----AspProGluGluIleAsnSerSerAlaProArgAsp 202
Db 1451 AATCCCACTTTCAGTGTGTAGACGAATCCAGAGCTTTCAACAAGACTCAGATCAGAGAC 1510
Qy 203 AspAspProGluAsnArgSerLysPheTyrGluSerAlaSer----- 216
Db 1511 AAGAACGTCCTAGGAGCGGTAAATTTCTTCTGAGGACCCCAAAAGCACCACTTGTAAATA 1570
Qy 217 AlaArgLysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThr 236
Db 1571 TCCCAAGGCGCGAGTAAACGAGAAGAGAGTCCCTTCTCAGAAAGCAATGGAG 1630
Qy 237 PheGluProThrAsnProPheArgValValLeuArgProSerTyrLeuTyrArgGly 256
Db 1631 TTCAAGTCGAAGCACCTTTCAACATGCAGATAATGATCGAGTCTTATGCTACCTGCGA 1690
Qy 257 CysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPhe 276
Db 1691 TTTTTCATGAACATCGCTCGGAGTTCGTCAAGGAGTCCCTCCCGCAGCAGCAAG--- 1747
Qy 277 IleLysValGlnLeuAlaGlu-----LysGlnTyrProValArgCysLeuTyrLys 293
Db 1748 ---AAGATGACGCTCTGGGACCCGAGTGGGAGCGCGTGGAGCTCAGTACGTCTACTAC 1804
Qy 294 AlaGlyArgAla-----LysPheSerGlnGlyTyrTyrGluPheThrLeuGluAsn 311
Db 1805 GCGAGACCGCGCGCGCTCTTCCGTGGCGGTGGGTAGTTCGCTCTGGGGAACAAC 1864
Qy 312 LeuGlyGluGlyAspValCysValPheGluLeuLeuArgThrArgAspPheValLeuLys 331
Db 1865 CTGGAAGAAGTCGAGCTCTGCGTCTTCGAGCTCTTCGAGGAGGAGCAAC-----ATGCGG 1918
Qy 332 ValThrAlaPheArgVal 337
Db 1919 GTGCACATATACAGGGTC 1936
```

## RESULT 14

```
US-10-437-963-89631/c
; Sequence 89631, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
```

APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 89631  
LENGTH: 1477  
TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_8836C.1  
US-10-437-963-89631

Alignment Scores:  
Pred. No.: 66-25 Length: 1477  
Score: 281.00 Matches: 104  
Percent Similarity: 42.93% Conservative: 60  
Best Local Similarity: 27.22% Mismatches: 152  
Query Match: 15.65% Indels: 66  
DB: 17 Gaps: 15

US-10-088-187A-11 (1-341) x US-10-437-963-89631 (1-1477)

QY 3 ArgProPheHisLeuLeuPheSerSerThrIleGluGlu---LysArgLeuArg 21  
Db 1447 AGAGCTCTCTCGTGCTTCTC-----ACCTGCATCTCTCGAGCGGATGAAA 1397

QY 22 ValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeuThr 41  
Db 1396 ATCCCATCATCTCAACCAATGCTTCAGAAATCAACCTACTGCGATGCTTCTCGTT 1337

QY 42 ValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrpPhe 61  
Db 1336 GATCGAAGTGGAAATAAATGAGTGCAGAGCTGACCTTCAGACTCTGAGGGATTCCTTT 1277

QY 62 GlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeuIle 81  
Db 1276 GTGCATGATGGAAGAAATTTGTACAGATACTCAATCACTGCTGGTCAATTTTAGTA 1217

QY 82 PheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSerGlu 101  
Db 1216 TTCACATTATGATAAACGGTCAACAATCTCAGTGACAGTCTTT-----GAACCTTCAGGC 1163

QY 102 IleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArg 121  
Db 1162 ATTGACAAGATATCAACT-----TTCTCTGCTCACCT-----TCCAAATAATGTG 1118

QY 122 LeuPheGluAspLeuGluAspGluAlaGluValIlePheProSerSerValTyrPro 141  
Db 1117 ATCATTAACACAAAGTGTATGAGGGGCGATGGTCCAGCTATCACACA----- 1064

QY 142 SerProLeuProGluSerThrValProAlaAsnLys-----GlyTyr 155  
Db 1063 -----GAGAAATGGTCCCGCCCTGAAGGAAATAATGAAATCAAGGAAAG 1016

QY 156 AlaSerSerAlaIleGlnThrLeuPheThrGlyProVal-----LysAlaGlu 171  
Db 1015 AGAACATAGAGATAGATATACCTATGGAAGATAGGTGGTGTATTTAAGAAATCATCT 956

QY 172 GluProThrProThrProLysLysArg-----GlyArgLys 186  
Db 955 GAGCTTAATGTTGCGAAAGTTCTAGAGAAACGTCGAGGAGCTTCTGCGCAAGTCC 896

QY 187 Lys-----LysAsnAlaAspProGluGluIle 195  
Db 187 Lys-----LysAsnAlaAspProGluGluIle 195

Db 895 AAGTAACATCGACAAAGCCATAACAGTACCAGAGGTTTCATCATGACGCTCAGATGAAGAT 836  
QY 196 AsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrGluSerAla 215  
Db 835 AATTCCTTAGC---AAGTCACCAATCCACCATTCCTGATGCGCTTTCTGTCTGTGAG 779  
QY 216 SerAlaArg----- 218  
Db 778 GTGTCCTCGGCTGGTAGATGTTTCAAGGGTCAAGACAACTAACAGTAATATCACAG 719  
QY 219 LysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAlaLysThrPheGlu 238  
Db 718 AGCGCTCCGGTAACTGAAGCAGAGAGGACCATCTCTTCAAGAGGCAAGAGATTAA 659  
QY 239 ProThrAsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIle 258  
Db 658 TCTAAGAATCTCTTTCGCGTGCAAAATAATGATGAGTCAATATGCTATGTTGATTTTC 599  
QY 259 MetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLys 278  
Db 598 ATGAATATCCCATGTGAATTTTCCGTGAGTGTCTTCCCATACCAACAAAGGATCACA 539  
QY 279 ValGlnLeuAlaGlu---LysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAla 297  
Db 538 CTCTGGGATCCACAGGAGGCTTGGGAAGTTAACTATGCTACTACAGTATCGTTCT 479  
QY 298 -----LysPheSerGlnGlyTyrTyrGluPheThrLeuGluAsnLeuGlyGluGly 315  
Db 478 GTTGGTCTTTCAGTGTGCTGGGCAAAATTTGCTTGGCAACAATCTGGAGAAGTTT 419  
QY 316 AspValCysValPheGluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPhe 335  
Db 418 GATGTGTGTCTTCGAGCTTGTCCAGAGGATAAC-----ATAAAGTGCACATCTAC 365  
QY 336 ArgVal 337  
Db 364 AGAGTT 359

RESULT 15  
US-10-425-114-32044  
; Sequence 32044, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 32044  
; LENGTH: 1455  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLB73255A02\_FLI  
US-10-425-114-32044

Alignment Scores:  
Pred. No.: 7,86e-25 Length: 1455  
Score: 280.00 Matches: 89  
Percent Similarity: 44.55% Conservative: 54  
Best Local Similarity: 27.73% Mismatches: 136  
Query Match: 15.59% Indels: 42  
DB: 16 Gaps: 13

US-10-088-187A-11 (1-341) x US-10-425-114-32044 (1-1455)



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 30, 2004, 06:15:53 ; Search time 6749 Seconds  
(without alignments)  
1841.154 Million cell updates/sec

Title: US-10-088-187A-11

Perfect score: 1796  
Sequence: 1 MRPFPFKLIFSSTQEKRL.....LLRTDRFVLKVTAFRNVYV 341

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlh  
-O=/csp2\_1/USPTO spoal/US10088187/runat\_29122004\_115842\_127/app\_query.fasta\_1.519  
-DB=EST\_QFMT=fastap SUPFIX=first -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.coi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10088187 @CNC 1 1 3437 @runat\_29122004\_115842\_127 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hcc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gest1:\*  
9: gb\_gest2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1789	99.6	1307	CNS0A70B	BX822585 Arabidops
2	1769	98.5	1441	CNS0A6ON	BX823300 Arabidops
3	1717	95.6	1375	CNS0A6SO	BX824523 Arabidops
4	1142.5	63.6	1346	CNS0A6P	BX816536 Arabidops
5	1141.5	63.6	1420	CNS0A03Z	BX816391 Arabidops
6	1141.5	63.6	1490	CNS0A086	BX815859 Arabidops
7	1061.5	59.1	880	CNS09469	BX841843 Arabidops
8	1016.5	56.6	780	CK120978	CK120978 204014.pl
9	934.5	52.0	906	CO108220	CO108220 GR_Eb003

10	932	51.9	701	6	CD822418	BN25_045C
11	860	47.9	1041	5	BX835682	BX835682 GR_Eb003
12	811	45.2	834	7	CO108208	CO108208 GR_Eb003
13	789.5	44.0	669	1	AV821829	AV821829
14	786	43.8	645	6	CD823066	CD823066 BN25_047I
15	771.5	43.0	834	7	CO094348	CO094348 GR_Ea16G
16	707.5	39.4	612	6	CD813712	CD813712 BN15_020L
17	682	38.0	589	1	AV540912	AV540912
18	667	37.1	591	1	AI996485	AI996485 701666887
19	660	36.7	785	8	BZ061379	BZ061379 lki10h12.
20	657	36.6	614	1	AV539303	AV539303
21	655.5	36.5	748	7	CO084070	CO084070 GR_Ea49H
22	652	36.3	841	7	CO132469	CO132469 GR_Eb45F
23	650	36.2	721	8	BH969696	BH969696 ocd94e04.
24	626.5	34.9	699	8	BZ464608	BZ464608 BONRX74TR
25	623.5	34.7	870	4	BM360659	BM360659 GA_Ea003
26	623	34.7	872	7	CO128357	CO128357 GR_Eb25B
27	620.5	34.5	591	6	CA781580	CA781580 024G12AF
28	620	34.5	809	7	CO099865	CO099865 GR_Ea24P
29	617	34.4	617	4	BM814806	BM814806 EST529200
30	603	33.6	631	1	AJ504308	AJ504308
31	597.5	33.3	712	7	CO132468	CO132468 GR_Eb45F
32	583.5	32.5	624	2	AW584452	AW584452 N210502e
33	579	32.2	636	6	CD038665	CD038665 UTPPT010
34	577.5	32.2	821	7	CO108207	CO108207 GR_Eb003
35	577.5	32.2	822	7	CO108219	CO108219 GR_Eb003
36	576	32.1	726	5	BX835627	BX835627
37	575	32.0	453	2	AW705298	AW705298 EK59C02.Y
38	564.5	31.4	754	6	CA990755	CA990755 EST644263
39	556.5	31.0	732	6	CD835772	CD835772 BN45_046J
40	553.5	30.8	537	4	BM143791	BM143791 saj48910.
41	543	30.2	317	1	AV526843	AV526843
42	542	30.2	471	1	AI992821	AI992821 701493711
43	535.5	29.8	448	7	CN493904	CN493904 Mdfw2004a
44	534.5	29.8	480	1	AL383040	AL383040 MCBC11D09
45	533.5	29.7	812	4	BM357983	BM357983 GH_Ea000

#### ALIGNMENTS

RESULT 1  
CNS0A70B 1307 bp mRNA linear HTC 06-FEB-2004  
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
DEFINITION GSLTFFB50ZB05 of flowers and buds of strain col-0 of Arabidopsis  
thaliana (thale cress).  
ACCESSION BX822585  
VERSION BX822585.1 GI:42465399  
KEYWORDS HTC; GSLT\_cDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 1307)  
AUTHORS Castellani, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,  
Menard, G., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,  
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.  
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1307)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
COMMENT The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen)  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

Schachter V., Weissenbach J., Salanoubat M.  
 URV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information center for Protein  
 Sequences). 5 prime and 3 prime are assembled with Phrap.  
 http://www.genoscope.cns.fr/externe/sequences/banque\_Projet\_EF/Full  
 length  
 http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.

## FEATURES

Location/Qualifiers

source

1..1307  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLFBS02E05"  
 /tissue\_type="Flowers and buds"  
 /plasmid="pCMVSPORT 6"  
 complement(1..1307)  
 /gene="At3g18990"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.53e-181 Length: 1307  
 Score: 1789.00 Matches: 340  
 Percent Similarity: 99.71% Conservative: 0  
 Best Local Similarity: 99.71% Mismatches: 1  
 Query Match: 99.61% Indels: 0  
 DB: 3 Gaps: 0

US-10-088-187A-11 (1-341) x CNS0A70B (1-1307)

QY 1 MetProArgProPheHisLysLeuLeuPheSerThrIleGlnGluLysArgLeu 20  
 |||||  
 Db 75 ATGCCACGCCCTTCTTCATAGTTGATTTCTCATCTATTCACATATTCAGAAAAAGCTGTG 134  
 QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40  
 |||||  
 Db 135 AGGGTCCACAGATAAGTTGTGAGTAAATTCAGAGTATGAGCTTTCGGTGTGCTGTCACCTC 194  
 QY 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60  
 |||||  
 Db 195 ACAGTACCTGTAGTGTGATGTTTGGCGGTGTAGGACTAAGGAAGCTGACAAACAAATTTGG 254  
 QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80  
 |||||  
 Db 255 TTTCAGATGTTTGGCAAGATTTGTTGACCTTACTCCATTCGATGGTATCTTTTG 314  
 QY 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100  
 |||||  
 Db 315 ATTTTATAGATATGAAGAACTCTGCTTTCAGCGTCTACATTTTCAATTTATCCACTCT 374  
 QY 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120  
 |||||  
 Db 375 GAGATCAATTTACATTCACCGCTCTCATGGATTCGCTCAACCACTTCAAACGCCCC 434  
 QY 121 ArgLeuPheGluAspLeuGluAspGluLeuValIlePheProSerSerValTyr 140  
 |||||  
 Db 435 CGTTTGTGAGACCTTGAAGATGAAGTCCGAGGTCACTTTCCTTCTCTGTGTAC 494  
 QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160  
 |||||  
 Db 495 CCATCACCATTCTCTGAGTCTACAGTACCAGCAACAAAGGGTATGCTAGTTCAGCCATC 554  
 QY 161 GlnThrLeuPheThrGlyProValLysAlaGluGluProThrProThrProLysIlePro 180  
 |||||  
 Db 555 CAAACCTTGTTCCTGAGACCTTAAAGCTGAAGGCCAACGCCCAACCCCAAAATACCT 614  
 QY 181 LysLysArgGlyArgLysLysAsnAlaAspProGluGluIleAsnSerSerAlaPro 200  
 |||||  
 Db 615 AAAAGAGAGGGAGGAGAGAGAAATGCTGATCTGAGGAAATAAACTCATCAGTCCG 674  
 QY 201 ArgAspAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220  
 |||||

Db 675 CGAGATGATGATCCAGACACCGTTCAAAGTTCTACGAGAGTGTCTTCGCCGAGAAAGAGA 734  
 QY 221 ThrValThrAlaGluGluArgGluArgAlaIleAlaIleAlaLysThrPheGluProThr 240  
 |||||  
 Db 735 ACCGTGACTCGAGAGAAAGAGAGAGCCATCAATGACGACCAACCGTTCGAACCAACA 794  
 QY 241 AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260  
 |||||  
 Db 795 AACCCCTTCTTCAGAGTGGTCTCGCACCATCTCTATACAGAGGTTCATCATGTAT 854  
 QY 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGln 280  
 |||||  
 Db 855 CTTCTCTCTGGGTTTGTCTGAGAAGTACCTAAGTGGATCTCCGGTTTCATCAAGTCCAG 914  
 QY 281 LeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSer 300  
 |||||  
 Db 915 CTTGCGGAGAAACAATGGCCTGTTCGATGTCTCTACAAAGCCGGAGAGCCAAATTCAGT 974  
 QY 301 GlnGlyTrpTyrGluPheThrLeuGluAsnLeuGlyGluGlyAspValCysValPhe 320  
 |||||  
 Db 975 CAAGATGTAGCAATTCATCTCTAGAACCACTTAGGAGAGAGAGAGCTCTGTGTGTTT 1034  
 QY 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyr 340  
 |||||  
 Db 1035 GAGCTGCCAGAACCCAGAGATTTGTTTGAAGTGAACAGCTTTCGAGTCAACGAGTAC 1094  
 QY 341 Val 341  
 |||||  
 Db 1095 GTC 1097  
 RESULT 2  
 CNS0A6ON 1441 bp mRNA linear HTC 06-FEB-2004  
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 DEFINITION Arabidopsis thaliana (thale cress).  
 ACCESSION BX823300  
 VERSION BX823300.1 GI:42464860  
 KEYWORDS HTC; GSLT cDNA  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 1441)  
 AUTHORS Castellani V., Aury J.M., Jaillon O., Wincker P., Clepet, C.,  
 Menard, M., Cruaud, C., Quetier, F., Scarpetti, C., Schachter, V.,  
 Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.  
 Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
 A Combined Approach to Evaluate and Improve Arabidopsis Genome  
 Annotation  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1441)  
 AUTHORS Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 Web : www.genoscope.cns.fr)  
 COMMENT The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castellani  
 V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,  
 Schachter V., Weissenbach J., Salanoubat M.  
 URV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information center for Protein  
 Sequences). 5 prime and 3 prime are assembled with Phrap.  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
 length  
 http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.  
 FEATURES Location/Qualifiers  
 source 1..1441

/organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLTSL202G12"  
 /tissue\_type="Adult vegetative tissue"  
 /plasmid="pCMVSPORT 6"  
 complement(1..1441)  
 /gene="At3g18990"

gene

## ORIGIN

## Alignment Scores:

Pred. No.: 2,51e-179 Length: 1441  
 Score: 1769.00 Matches: 340  
 Percent Similarity: 99.71% Conservative: 0  
 Best Local Similarity: 99.71% Mismatches: 1  
 Query Match: 98.50% Indels: 1  
 Ds: 3 Gaps: 0

US-10-088-187A-11 (1-341) x CNS0A60N (1-1441)

QY 1 MetProArgProPheHisLysLeuIlePheSerSerThrIleGlnGluLysArgLeu 20  
 Db ATGCCACGCCCTTCTTCATTAAGTTGATTTCTCATCTATCCAGAGAAAGCTGTG 291  
 QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40  
 Db AGGGTCCAGATAAAGTTGTGAGTAATTCAGAGATGAGCTTCGGTGTGCTGCTGCACTC 351  
 QY 41 ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysIleTrp 60  
 Db ACAGTACCTGATGGTCATGTTGGCGTGTAGACTAAGAAAGCTGCAACAAATTTGG 411  
 QY 61 PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeu 80  
 Db TTTCAAGATGGTTCGACAGAGTTTGTGACCGTTACTTCATTCGATGGTATGTTTGTG 471  
 QY 81 IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer 100  
 Db ATTTTATAGATATGAAGAAACTCTGCCCTTCAGCGTCTACATTTTCAATTTATCCACTCT 531  
 QY 101 GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla 120  
 Db GAGATCAATTAACCATTCACCGCTCTCATGGATTCGCTCAACCACTTCAAACGCCCC 591  
 QY 121 ArgLeuPheGluAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyr 140  
 Db CTTTGTGTTGAGACTTGAAGATGAAGATGCCGAGGTCATCTTCTCTCTCTGTGTAC 651  
 QY 141 ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIle 160  
 Db CCATCAACCATTCCTCGAGTCTACAGTACCAGCCCAAAAGGTATGCTAGTTCAGCCATC 711  
 QY 161 GlnThrLeuPheThrGlyProValLysAlaGluGluProThrProThrProLysIlePro 180  
 Db CAACCTTTGTTCTACGTGACCACTTAAGCTGAAGAGCCCAACGCCCAACCAAAATACCT 771  
 QY 181 LysLysArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaPro 200  
 Db AAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 830  
 QY 201 ArgAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg 220  
 Db CGAGATGATGATCCAGAGAACCGTTCAAAAGTTCTACGAGAGTCTTCGCGAGAAAGAGA 890  
 QY 221 ThrValThrAlaGluGluArgGluArgAlaIleAsnAlaLysThrPheGluProThr 240  
 Db ACCGTGACTGCAGAGAAAG 950  
 QY 241 AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr 260  
 Db AACCTTCTTCTCAGAGTGGTCTCGGACCATCTCTATCTATACAGAGGTTGTCATGATAT 1010

QY 261 LeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysValGln 280  
 Db CTTCTCTCTGGTTTGTCTGAGAAAGTACCTTAAGTGGGATCTCGGGTTTCATCAAGTCCAG 1070  
 QY 281 LeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSer 300  
 Db CTTGCGGAGAACATGCGCTTTCATGATGTTCTCTAAAGCCGCGAGAGCCAAATTCAGT 1130  
 QY 301 GlnGlyTyrTyrGluPheThrLeuGluAsnLeuLeuGlyGluGlyAspValCysValPhe 320  
 Db CAAGGATGTTAGCAATTCCTCTAGAGAACAACTTAGGAGAGGAGAGCGTCTGTGTGTTT 1190  
 QY 321 GluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgValAsnGluTyr 340  
 Db GAGCTGCTCAGAACCCAGAGATTTCTGTTTGAAGTGCAGAGCCTTTTCGAGTCAACGAGTAC 1250  
 QY 341 Val 341  
 Db 1251 GTC 1253

RESULT 3  
 CNS0A60S  
 LOCUS  
 DEFINITION  
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSUTPGH542B04 of Hormone Treated Callus of strain col-0 of Arabidopsis thaliana (thale cress).  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Arabidopsis thaliana  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 1375)  
 Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Searpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
 Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
 Unpublished  
 2 (bases 1 to 1375)  
 Genoscope.  
 Direct Submission  
 Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
 The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.  
 URV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full length  
 http://www.genoscope.cns.fr/cgi-bin/ggb7?source=Arabidopsis.  
 Location/Qualifiers  
 1. 1375  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLTPGH542B04"  
 /tissue\_type="Hormone Treated Callus"  
 /plasmid="pCMVSPORT 6"  
 complement(1..1375)  
 /gene="At3g18990"

FEATURES  
 source  
 gene  
 ORIGIN

Db	1125	CAAGGATGGTACGAATTCCTCTAGAGAACCACTTAGGAGAGGAGACGCTGTGTGTCT	1184
Qy	321	GlulLeuLeuArgThrArgAspPheValLeuLeuValThrAlaPheArgValAlaGluTyr	340
Db	1185	GAGCTGCTCAGAACCCAGAGATTTCGTTTGAAGAGTGCACGCTCTCGAGTCAACGAGTAC	1244
Qy	341	Val	341
Db	1245	GTC	1247
RESULT 4			
CNSOAD6P			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
gene			
ORIGIN			
Alignment Scores:			
Pred. No.:	6, 61e-112	Length:	1346
Score:	1142.50	Matches:	243
Percent Similarity:	76.22%	Conservative:	23
Best Local Similarity:	69.63%	Mismatches:	27
Query Match:	63.61%	Indels:	57
DB:	3	Gaps:	10
US-10-088-187A-11 (1-341) x CNSOAD6P (1-1346)			
Qy	1	MetProArgProPhePheHisLysLeuLeuPheSerSerThrLeuGlnLysArgLeu	20
Db	225	ATGCCACGCCCTTTCTTCATGAGTTGATTTCTCATCTCATCTCCATCCAGAGGGAGCTG	284
Qy	21	ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu	40
Db	285	ATGGTCCACAGATAAGTTTGTAGTGAATTCGAGGATGAGCTTCGGTTGCTGTTCACATC	344
Qy	41	ThrValProAspGlyHisValTrpArgValGlyLeuArgLysAlaAspAsnLysLeuTrp	60
Db	345	ACAGTACCTGATGGTCATGTTTGGCGGTAGGACTAAGGAGAGCTGACCAAAATTTGG	404
Qy	61	PheGlnAspGlyTrpGlnGluPheValAspArgTyrSerIleArgIleGlyLysLeuLeu	80
Db	405	TTTCAAGATGGTGGCAAGAGTTTGTGACCGTTACTTCATTCGATTTGTTGTTTGG	464
Qy	81	IlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSer	100
Db	465	ATTTTATAGATGAAGAAATCTCGCTTACGGTCTACATTTTCGATTTATCCCACTCT	524
Qy	101	GluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAla	120
Db	525	GAGATCAATATACATTCACCGGTCTCATGGATTCGCTCAGCACCATTTCAAACGCCCC	584
Qy	121	ArgLeuPheGluAspLeuGluAspGluAlaGluValIlePheProSerSerValTyr	140
Db	585	CGTTGTTTGAAGACCTTGAAGATGAAGATCCGAGGTCATCTTCTCTCTGTATC	644
Qy	141	ProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaLe	160
Db	645	CCATCACCATCTCTGAGTCTACAGTACCAAGCAAAAGGATATGCTAGTTTCAGCCATC	704
Qy	161	GlnThrLeuPheThrGlyProValLysAlaGluGluProThrProThrProLysIlePro	180
Db	705	CAAACTTGTCTACTGGACAGTTAAAGCTGAGAGCCCAACCCCAACCCCAAAATACCT	764
Qy	181	LysLysArgGlyArgLysLysLysAsnAlaAspProGluGluIleAsnSerSerAlaPro	200
Db	765	AAAAAGAGAGGGAGGAAGAAATAATGCTGATCTCTGAGGAAATAAACTCATCAACTCCG	824
Qy	201	ArgAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArg	220
Db	825	CGAGATGATGATCCAGAGAACCGTTCAAAAGTTCTAGAGAGTGTCTCGAGAAAGAGA	884
Qy	221	ThrValThrAlaGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThr	240
Db	885	ACCGTACCTGCAGAAAGAGAGAGAGCCATCATCGACCAAAACGTTTCGACCAACA	944
Qy	241	AsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyr	260
Db	945	AACCCCTTCTCAGAGTGGTCTCGACCATCTATCTATACAGAGTGTGATCATGTAT	1004
Qy	261	LeuProSerGlyPheAlaGluLysTyrLeuSerGlyLeuSerGlyPheIleLysValGln	280
Db	1005	CTTCTCTCTGGGTTTCTGAGAAAGTACCTAAGTGGGATCTCCGGGTTTCAAAAGTCCAG	1064
Qy	281	LeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSer	300
Db	1065	CTCGCGGAGAAACAATGGCTGTTCGATGTTCTCTACAAAGCCGGGAGAGCCAACTCACT	1124
Qy	301	GlnGlyTrpTyrGluPheThrLeuGluAsnAsnLeuGlyLysValCysValPhe	320





```
QY 99 sSerGluIleAsnTyrHisSerThrGly-----LeuMetAspSerAlaHisAs 115
Db 351 CTCTGAGATCAACACCACCTTAGTAGTGAAGCTCTTATGCAAAATGGATTCGCCAGAA 410
QY 115 nHisPhe---LysArgAlaArgLysPheGluAspGluAspGluValIle 134
Db 411 TCAGTTTCAACAACGCGTGGATTTGTTGAAGATCTGGAATCAAGATGCGAAGTCTAT 470
QY 134 ePheProSerSerValTyrProSerProLeuProGluSerThrValProAlaAsnLysG 154
Db 471 TTATCCATCGAAC-----CCTGAATCTACTGAACCAAGTGAATAAGG 512
QY 154 YTYR---AlaSerSerAlaIleGlnThrLeuPheThrGlyProValLysAlaGluPro 173
Db 513 TTATGCGCGTCTTACAGCCATCCAAAGCTTTTTCAAA---GAATCTAAAGCTGAAGAA-- 567
QY 173 oThrProThrProLysIleProLysArgLysGlyArgLysLysAsnAlaAspProG 193
Db 568 -----ACGCCAAAGGTACTTAAGAAGAGAGAGAGAGAGAGAGAGAAATCCTAATCCCGA 620
QY 193 uGluIleAsnSerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrG 213
Db 621 GGAAGTAAACTTCTCAACTCCCGTGGAGATGACTCAGAGAACCGCTCAAGTCTACGA 680
QY 213 uSerAlaSerAlaArgLysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAl 233
Db 681 GAGTGTCTCTGCTAGAAAGAGAACTGTAATCTCAGAGAGAAAGAGAGAGAGCGTCAATGC 740
QY 233 alalysThrPheGluProThrAsnProPhePheArgValValLeuArgProSerTyrI 253
Db 741 AGCCAAACATTCGAAACCAACAACTTACTTACTTAAAGTGTGTTGCGGACCATCATCT 800
QY 253 uTyrArgGlyCysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyI 273
Db 801 ATACAGAGGTTCATCACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 860
QY 273 eSerGlyPheIleLysValGlnLeuAlaGluLysGlnTrpProValArgCysLeuTyrI 293
Db 861 ATCTGTTTTCATCAAGCTCCAGCTCGGTGAGAAACAAATGCGGAGTGGTGGTGGTGG 920
QY 293 salaglyArgAlaLysPheSerGlnGlyTyrTrpTyrGluPheThrLeuGluAsnLeuG 313
Db 921 AGCAGGAGAGTAAAGTTTAGCCACAGGATGGTATGAGTTTCACTACACGAGAAATATAG 980
QY 313 yGluGlyAspValCysValPheGluLeuLeuArgThrArgAspPheValLeuLysValTh 333
Db 981 CGAAGAGATGATGTGTGTTTGAGCTACTCAGAACTCGGATTTCTGCTGGAAGTCAAC 1040
QY 333 rAlaPheArgValAsnGluTyrVal 341
Db 1041 CGCCTTCGTGTCATGAGTATGTG 1065

RESULT 7
CNS09Y69
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTSL76ZC03 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION
BX841843
VERSION
BX841843.1 GI:42454500
KEYWORDS
HTC; GSLT cDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
1 (bases 1 to 880)
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome

Annotation
Unpublished
2 (bases 1 to 880)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_BF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
Location/Qualifiers
1..880
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strains="Col-0"
/db_xref="taxon:3702"
/clone="GSLTSL76ZC03"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
complement(1..880)
/gene="At1g49480"

Alignment Scores:
Pred. No.: 1.9e-103 Length: 880
Score: 1061.50 Matches: 226
Percent Similarity: 83.67% Conservative: 25
Best Local Similarity: 75.33% Mismatches: 22
Query Match: 59.10% Indels: 28
DB: 3 Gaps: 9

US-10-088-187A-11 (1-341) x CNS09Y69 (1-880)
QY 51 GlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGlyTrpGlnGluPheValAsp 70
Db 1 GGACTAAGAAAGCCACACAAATCTGTTTCAAGACGGTTGCGAGGAGTTGTTCAC 60
QY 71 ArgTyrSerIleArgIleGlyTyrLeuLeuIlePheArgTyrGluGlyAsnSerAlaPhe 90
Db 61 GCTTCTCCATTCGGATTGGT-----TTCAGATACAA-----GTT 96
QY 91 SerValTyrIlePheAsnLeu---Ser-HisSerGluIleAsnTyrHisSerThrGly-- 108
Db 97 ACAGTCTACATTTTCAATTTATCTCTCCACACTCTGAGATCAACCACTTCTAGTAGTA 156
QY 109 -----LeuMetAspSerAlaHisAsnHisPhe---LysArgAlaArgLeuPheG 124
Db 157 AGCTCTTATGCAATGGATTTCGACAGATCAGTTCACAAACGCTGCGATTGTTCGA 216
QY 124 uAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyrProSerProle 144
Db 217 AGATCCTGAACCTCAAGATGCTAAGGTCAATTTATCCTCAAC----- 259
QY 144 uProGluSerThrValProAlaAsnLysGlyTyr---AlaSerSerAlaIleGlnThrle 163
Db 260 -CCTGAATCTACTGAACCAAGTGAAGGTTATGCGGTTTCTACAGCCATCCAAAGCTT 318
QY 163 uPheThrGlyProValLysAlaGluGluProThrProThrProLysIleProLysIle 183
Db 319 TTTTCAA---GAATCTAAAGCTGAAGG-----ACGCCAAAGGTACTTTAAGAGAG 366
QY 183 gGlyArgLysLysLysAsnAlaAspProGluIleAsnSerSerAlaProArgAspAs 203
```



QY 306 eThrLeuGluAAsnLeuGlyCluGlyAspValCysValPheGluLeuLeu 323  
 |||||  
 Db 728 CACACTCGAACAATATATAGCGAGAGATGATGTGTGTGAGCTACT 779

RESULT 9  
 CO108220 GR\_Eb0039N17.r GR\_Eb Gossypium raimondii cDNA clone GR\_Eb0039N17  
 LOCUS 3', mRNA sequence. 906 bp mRNA linear EST 16-JUN-2004  
 DEFINITION  
 ACCESSION CO108220.1 GI:48806906  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Gossypium raimondii  
 Gossypium raimondii  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
 1 (bases 1 to 906)  
 Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,  
 Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and  
 Wing,R.A.  
 Global assembly of Cotton ESTs  
 Unpublished (2004)  
 Contact: Rod A. Wing  
 Arizona Genomics Institute  
 The University of Arizona  
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA  
 Tel: 520 626 9595  
 Fax: 520 621 1259  
 Email: http://genome.arizona.edu  
 Plate: 0039 row: N column: 17.

FEATURES  
 source  
 1..906  
 /organism="Gossypium raimondii"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:29730"  
 /clone="GR\_Eb0039N17"  
 /tissue\_type="floral"  
 /dev\_stage="3 to +3 DPA"  
 /lab\_host="DH10B"  
 /clone\_lib="GR\_Eb"  
 /note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:  
 EcoRI; Library made by Invitrogen with RNA supplied by  
 Wendle lab. Directional cloned into NotI-EV. Colonies  
 plated/picked by AGI. More glycerol clones held in -80."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 9,39e-90 Length: 906  
 Score: 934.50 Matches: 189  
 Percent Similarity: 75.24% Conservative: 45  
 Best Local Similarity: 60.77% Mismatches: 54  
 Query Match: 52.03% Indels: 23  
 DB: 7 Gaps: 7

US-10-088-187A-11 (1-341) x CO108220 (1-906)  
 QY 9 LeuilePheSerThrleGlnLysArgLeuArgValProAspLysPheValSer 28  
 |||||  
 Db 2 CTTATTCTCTTACCACCTCCCAAGAAAGAACTGAGTCCCTGATACTTCGTTAAG 61  
 |||||  
 QY 29 LysPheLysAspGluSerValalaValalaLeuThrValProAspGlyHisValTrp 48  
 |||||  
 Db 62 AAATTCAGGAGTGAATCTTCTGTGTCGCGCTCTCACTGTTCCTGAGCGTATGTTGG 121  
 |||||  
 QY 49 ArgValGlyLeuArgLysAlaAspAsnLysileTrpPheGlnAspGlyTrpGlnGluPhe 68  
 |||||  
 Db 122 CGTAGGATGAATAAGAAAGTTCACACAGGTTTGGTTTCAGAGAGTTGGCAGAGTTT 181  
 |||||  
 QY 69 ValAspArgTyrSerileArgleGlyTyrLeuLeuilePheArgTyrGluGlyAsnSer 88  
 |||||  
 Db 182 CTAGAGCGTATTATATTCTGTGCTAGCTACGTATTGTTGTTTATAGATACGAGAAATCT 241  
 |||||

QY 89 AlapheSerValTyrilePheAsnLeuSerHisSerGluileAsnTyrHisSerThrGly 108  
 |||||  
 Db 242 GCTTTTCAGTGTAGTATATTTAATTGTACAACTCGAAATAAATACTATCAGACTAATGCC 301  
 |||||  
 QY 109 LeuMetAspSerAlaHisAsnHisPheLysArgAlaArgLeuPheGluAspLeuGluAsp 128  
 |||||  
 Db 302 CTCGTTGGTACTCAATACAACTCACGAAACAAATATCCA---TTTGAACAACCTTGAAGAT 358  
 |||||  
 QY 129 GluAspAla-----GluValilePheProSerSerValTyrProSerPro 143  
 |||||  
 Db 359 GATGAATGCATCTCTCCAGCACTTCAGAAATTTGTTGGTCTTAAACTTAACTCAACTGC 418  
 |||||  
 QY 144 Leu-----ProGluSerThrValProAlaAsnLysGlyTyrZalaSerSerAlaile 160  
 |||||  
 Db 419 ATAAACTGGGGTGGTGGTATCCCAATCTTCAAACTTCAAGGGTGTGAATAATCAACTATT 478  
 |||||  
 QY 161 Gln-----ThrLeuPheThrGlyProValLysAlaGluGluProThrProThr 176  
 |||||  
 Db 479 CGAGTCAAAATTCATACACTTCAGGTTTCAGGTGCAATGAAACCAAGAA----- 523  
 |||||  
 QY 177 ProLysileProLysArgGlyArgLysLysAsnAlaAspProGluGluileAsn 196  
 |||||  
 Db 524 -----CCAAAAGCGTGGAGAAAGCCGAAAG---TTTGATCTTAACTGCGAGGAT 571  
 |||||  
 QY 197 SerSerAlaProArgAspAspProGluAsnArgSerLysPheTyrGluSerAlaSer 216  
 |||||  
 Db 572 TCATCTGCTGGACGTTGAAGATGATGTTGATATGCGCTTTAGATGTTATGAAGATGCTTCA 631  
 |||||  
 QY 217 AlaArgLysArgThrValThraAlaGluArgGluArgAlaileAsnAlaLysThr 236  
 |||||  
 Db 632 GCCCGAAAGAGAACCGTGCACGCTGAAGAAAGAGAGAGCAATCAATGTCAGCAAGCA 691  
 |||||  
 QY 237 PheGluProThrAsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGly 256  
 |||||  
 Db 692 TTTGAGCCCAACTAACCTTCTGCGAGGGTGGCTTGGCGACCACTATCTATCTGACAGGGGA 751  
 |||||  
 QY 257 CysileMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyLysileSerGlyPhe 276  
 |||||  
 Db 752 TGTATTATGTACTTACCGTGGCTTGTGCTGAGAGACATCAAGTGGGGTTCCTGATCC 811  
 |||||  
 QY 277 IleLysValGlnLeuAlaGlu---LysGlnTrpProValArgCysLeuTyrLysAlaGly 295  
 |||||  
 Db 812 ATTAACCTTCAGCTTCCTGATGGGAGACAGTGGTCTGTAAGATGCTTTTATAAGAGGAGC 871  
 |||||  
 QY 296 ArgAlaLysPheSerGlnGlyTrpTyrGluPhe 306  
 |||||  
 Db 872 AAGCTAANTTCAGTCAGGATGGTATGAATTT 904  
 |||||

RESULT 10  
 CD822418 BN25.045C06F020108 BN25 Brassica napus cDNA clone BN25045C06, mRNA  
 LOCUS 701 bp mRNA linear EST 10-JUL-2003  
 DEFINITION  
 ACCESSION CD822418  
 VERSION CD822418.1 GI:32504358  
 KEYWORDS EST.  
 SOURCE Brassica napus (rape)  
 ORGANISM Brassica napus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 701)  
 Genoplatte.  
 Genoplatte, a major partnership french program in plant genomics  
 Unpublished (2003)  
 JOURNAL  
 COMMENT  
 Contact: Genoplatte  
 Genoplatte  
 93, rue Henri Rochefort 91025 EVRY CEDEX France  
 Tel: 33 1 69 47 54 00  
 Fax: 33 1 69 47 54 10  
 This sequence has been generated in the framework of the french  
 plant genomics programme 'Genoplatte' (http://www.genoplatte.com)

```
and http://genoplante-info.infobiogen.fr/.
FEATURES
  source
    1..701
      Location/Qualifiers
        organism="Brassica napus"
        mol_type="mRNA"
        cultivar="Jet Neuf"
        db_xref="taxon:3708"
        clone="BN25045C06"
        tissue_type="seed"
        clone_lib="BN25"

ORIGIN
Alignment Scores:
Pred. No.: 1..21e-89 Length: 701
Score: 932.00 Matches: 180
Percent Similarity: 92.12% Conservative: 7
Best Local Similarity: 88.67% Mismatches: 14
Query Match: 51.89% Indels: 2
DB: 6 Gaps: 2

US-10-088-187A-11 (1-341) x CD822418 (1-701)

QY 1 MetProArgProPhePheHisLysLeuLePheSerSerThrIleGlnGluLysArgLeu 20
DB 92 ATGCCACGCCCTTTCTCCACAGCTGATTTCTCATCATCATATCCAGAAAAACGTCGT 151
QY 21 ArgValProAspLysPheValSerLysPheLysAspGluLeuSerValAlaValAlaLeu 40
DB 152 AGAGTTCCTGATAAGTTGTGAGTAATTCAGGACGAGCATCGGTTGCCGTTGCTCTC 211
QY 41 ThrValProAspGlyHisValTyrArgValGlyLeuArgLysAlaAsp---AsnLysIle 59
DB 212 ACAGTACCTGATGTCATGTTTGGGTGTAGGACTAAGGAAAGCTGACACACAAATC 271
QY 60 TrpPheGlnAspGlyTyrGlnGluPheValAspArgTyrSerIleArgIleGlyTyrLeu 79
DB 272 TGGTTTCAAGATGTTGGCAAGAGTTGTGTACCGTTACTCAATCCGATTTGGTTACCTT 331
QY 80 LeuIlePheArgTyrGluGlyAsnSerAlaPheSerValTyrIlePheAsnLeuSerHis 99
DB 332 TTGATATTAGATACGAAGGCAACTTCCTCTTCAGCGTCTACATTTACAATTTATCACAC 391
QY 100 SerGluIleAsnTyrHisSerThrGlyLeuMetAspSerAlaHisAsnHisPheLysArg 119
DB 392 TCCGCGATCACTACCACTCCACCGGTCTCATGGATTCGCACACACCACTTCAAGCT 451
QY 120 AlaArgLeuPheGluAspLeuGluAspGluAlaGluValIlePheProSerSerVal 139
DB 452 GCGCGTTTGTGTAAGACCCCGAAGATGTTGAGGTTGTGAGGTTGTTCACCCCTTCTCCTTTG 511
QY 140 TyrProSerProLeuProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAla 159
DB 512 TACCATCACAGCATCTGAGACTACTCGGCACCGCTAATAAGGCGATAGTTTCAGCT 571
QY 160 IleGlnThrLeuPheThr---GlyProValLysAlaGluGluProThrProThrProLys 178
DB 572 ATCCAGAGCTTCTTCGCTGAACCACTGTATAAGCTGAAGACACCAACCCCAAAA 631
QY 179 IleProLysLysArgGlyArgLysLysLysAlaAspProGluGluIleAsnSerSer 198
DB 632 GTTCTTAAAGAGAGAGGAGGAGGAAGAAGACGCTGATCTCTCGAGGAATAAATTCATCT 691
QY 199 AlaProArg 201
DB 692 GCTCCACGA 700

RESULT 11
LOCUS BX835682/c
DEFINITION BX835682 Arabidopsis thaliana Adult vegetative tissue Col-0
ACCESSION Arabidopsis thaliana cDNA clone GSLTUS48ZH11 3PRIM, mRNA sequence.
VERSION BX835682.1 GI:42529765

KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 1041)
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished (2004)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences).
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES
  source
    1..1041
      Location/Qualifiers
        organism="Arabidopsis thaliana"
        mol_type="mRNA"
        ecotype="Col-0"
        db_xref="taxon:3702"
        clones="GSLTUS48ZH11"
        tissue_type="Adult vegetative tissue"
        clone_lib="Arabidopsis thaliana Adult vegetative tissue
Col-0"

ORIGIN
Alignment Scores:
Pred. No.: 1..21e-81 Length: 1041
Score: 860.00 Matches: 177
Percent Similarity: 79.10% Conservative: 16
Best Local Similarity: 72.54% Mismatches: 31
Query Match: 47.88% Indels: 20
DB: 5 Gaps: 6

US-10-088-187A-11 (1-341) x BX835682 (1-1041)

QY 112 SerAlaHisAsnHisPheLysArg-----AlaArgLeu 122
DB 1009 ACTCGCACACACCACTTACTCGTCGACGCACTTTGCTTGGTTTACCGCACAGTA 950
QY 123 Phe-----GluAspLeuGluAspAlaGluValIlePheProSerSer 138
DB 949 CTGTACTACTAACGTACACGTTTGTCTGACGACACTGACACACGCTT-----ACTAGG 896
QY 139 ValTyrProSerProLeuProGluSerThrValProAlaAsnLysGlyTyr---AlaSer 157
DB 895 GACATTTAAACACACGAAACTGAACTACTGAACCAAGTAATAAGGTTATGGCGGTTCT 836
QY 158 SerAlaIleGlnThrLeuPheThrGlyProValLysAlaGluGluProThrProThrPro 177
DB 835 ACAGCCATCCAAAGCTTTTTCAAA---GAATCTAAAGCTGAAGAA-----ACGCC 788
QY 178 LysIleProLysLysArgGlyArgLysLysAsnAlaAspProGluGluIleAsnSer 197
DB 787 AAGGTACTTAAGACAGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 728
QY 198 SerAlaProArgAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAla 217
```

```

Db      727  TCAACACCCGGTGGAGATGACTCAGAGAACCCGCTCAAGTTCTACGAGAGTCTTCTGCT 668
Qy      218  ArgLysArgThrValThrAlaGluArgGluArgAlaAlaAsnAlaAlaLysThrPhe 237
Db      667  AGAAGAGAACTGTAACTCAGAGGAAAGAGAGAGCCGCTCAATGCAGCCAAACATTC 608
Qy      238  GluProThrAsnProPheArgValValLeuArgProSerTyrLeuTyrArgGlyCys 257
Db      607  GAACCAACAACATCTTACTTTAGAGTTGTCTCGCAGCATCATATCATACAGAGTTGC 548
Qy      258  IleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIle 277
Db      547  ATCATGTACTTGCATCTGGTTCCTGAGAAATACCTAAGTGGGATATCTGGTTTCATC 488
Qy      278  LysValGlnLeuAlaGluLysGlnTrpProValArgCysLeuTyrLysAlaGlyArgAla 297
Db      487  AGCTCCAGCTCGGTGAGAAACATGCGCCAGTGAGTGCTCTACAGCAGGAGAGCT 428
Qy      298  LysPheSerGlnGlyTyrTrpTyrGluPheThrArgAspPheValLeuLysValThrAlaPheArgVal 317
Db      427  AAGTTTAGCAAGATGGTATGAGTTCACTCGAGAACATATATAGGCGAAGGAGATGTA 368
Qy      318  CysValPheGluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgVal 337
Db      367  TGTGTGTTTGAGCTACTCAGAACTCGGATTTCTGTCGAAAGTCACCGCTTTCTGTGTC 308
Qy      338  AsnGluTyrVal 341
Db      307  AATGAGTATGTC 296

```

```

RESULT 12
LOCUS   CO108208
DEFINITION GR_Eb0039N11.r GR_Eb Gossypium raimondii cDNA clone GR_EB0039N11
3', mRNA sequence.
ACCESSION CO108208
VERSION   CO108208.1 GI:48806894
KEYWORDS EST.
SOURCE   Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 834)
Kim.H., Yu.Y., Kudrna,D., Hatfield,J., Strum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 0039 row: N column: 11.
FEATURES             location/Qualifiers
     1..834
     /organism="Gossypium raimondii"
     /mol_type="mRNA"
     /db_xref="taxon:29730"
     /clone="GR_Eb0039N11"
     /tissue_type="Floral"
     /dev_stage="3 to +3 DPA"
     /lab_host="DH10B"
     /clone_lib="GR_Eb"
     /note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Clones
plated/picked by AGI. More glycerol clones held in -80."
ORIGIN
AV821829

```

```

Alignment Scores:
Pred. No.:      1.66e-76      Length:      834
Score:          811.00      Matches:    169
Percent Similarity: 73.96%      Conservative: 44
Best Local Similarity: 58.68%      Mismatches: 52
Query Match:    45.16%      Indels:    23
DB:             7      Gaps:      6
US-10-088-187A-11 (1-341) x CO108208 (1-834)

Qy      9  LeuIlePheSerSerThrIleGln-GluLysArgLeuArgValProAspLysPheValSe 28
Db      2  CTTATTCTCTCTACCACTCCCAAGCAATGAATGAGATCCCTGATTAACCTTCGTAA 61
Qy      28  rlyPheLysAspGluLeuSerValAlaValAlaLeuThrValProAspGlyHisValTr 48
Db      62  GAAATTCAGGGATGAACCTTCTGTGCTGCCCTCTCACTGTCTGACGGTCACTGTTG 121
Qy      48  pArgValGlyLeuArgLysAlaAspAsnLysIleTrpPheGlnAspGlyTrpGlnGluPh 68
Db      122  CGGTGTAGGAATAAAGAAAGTTGACACACAGGTTTGGTTTCAGGAAGGTTGCAGAGTT 181
Qy      68  eValAspArgTyrSerIleArgIleGlyTyrLeuLeuIlePheArgTyrGluGlyAsnSe 88
Db      182  TCTAGAGCGTTACTATATTCTGTGTGGCTACGTATTGTTTGTAGTACGAGGAATTC 241
Qy      88  rAlaPheSerValTyrIlePheAsnLeuSerHisSerGluIleAsnTyrHisSerThrCl 108
Db      242  TGCITTTTCAGTGTAGTATATTAAATTTGTACAACTCGGAATAATACTATCAGACTAATGC 301
Qy      108  yLeuMetAspSerAlaHisAsnHisPheLysArgAlaArgLeuPheGluAspLeuGluAs 128
Db      302  CCTCGTTGGTACTCAATACATCAGGAACAAATATCCA---TTTGAACAACCTTGAGA 358
Qy      128  pGluAspAla-----GluValIlePheProSerSerValTyrProSerPr 143
Db      359  TGATGATGATCATCTCTCCAGCACTTCAGAAATTTGTTGTGGTCTAAACTTAAACAACATG 418
Qy      143  oLeu-----ProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAlaAl 160
Db      419  CATAACTGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 478
Qy      160  eGln-----ThrLeuPheThrGlyProValLysAlaGluGluProThrProth 176
Db      479  TCGAGTCAAAATGCACTACTTCAGGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 524
Qy      176  rProLysIleProLysLysArgGlyArgLysLysLysAsnAlaAspProGluGluLys 196
Db      525  -----CCAAAAAAGCGTGGGAGAAAGCGGAAG---TTTGATCCTTAACGTGAGGA 571
Qy      196  nSerSerAlaProArgAspAspAspProGluAsnArgSerLysPheTyrGluSerAlase 216
Db      572  TTCTATCTCTCGACGTGAAGATGATGTTGATATGCGCTTTAGATGTTTATGAAGTGGCTTC 631
Qy      216  rAlaArgLysArgThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysTh 236
Db      632  AGCCGGAAGAGAACCCGTGACAGCTGAAGAAAGAGAGAGAGCAATCATATGAGCCAAAGC 691
Qy      236  rPheGluProThrAsnProPheArgValValLeuArgProSerTyrLeuTyrArgGl 256
Db      692  ATTTGAGCCAACTAACCCCTTTCTGCAGGGTCTGCTTCGACCATCTTATCTGTACAGGGG 751
Qy      256  yCysIleMetTyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPh 276
Db      752  ATGTATTATGACTTACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 811
Qy      276  eIleLysValGlnLeuAlaGlu 283
Db      812  CATTAAACTTCAGCTTCTCTGAT 833
RESULT 13
AV821829

```

LOCUS AV821829 669 bp mRNA linear EST 01-APR-2002  
 DEFINITION AV821829 RAFL4 Arabidopsis thaliana cDNA clone RAFL04-16-P22 5', mRNA sequence.  
 ACCESSION AV821829  
 VERSION AV821829.1 GI:19863860  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 669)  
 AUTHORS Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.  
 TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Motoaki Seki  
 Plant Functional Genomics Research Group  
 RIKEN Genomic Sciences Center  
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
 Tel: 81-298-36-4359  
 Fax: 81-298-36-9060  
 Email: mseki@tc.riken.go.jp  
 An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified plasmid vector as a SstI/XhoI insert. Please visit our web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further details.  
 FEATURES  
 source  
 1. 669  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="RAFL04-16-P22"  
 /dev\_stage="rosette plants"  
 /lab\_host="SOLR"  
 /clone\_lib="RAFL4"  
 /note="Site 1: SstI; Site 2: XhoI; subjected to cold-treated(1,2,5,10,24 hr)"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2 5e-74 Length: 669  
 Score: 789.50 Matches: 171  
 Percent Similarity: 79.75% Conservative: 18  
 Best Local Similarity: 72.15% Mismatches: 22  
 Query Match: 43.96% Indels: 26  
 DB: 1 Gaps: 9  
 US-10-088-187a-11 (1-341) x AV821829 (1-669)  
 QY 51 GlyLeuArgLYeAlaAspAsnLysIleTrpPheGlnAspGlyTrpGlnLpPheValasp 70  
 DB 15 GCATTAAGCAAGCAACCAACAAATCTGTTTCAAGACGGTTGCGAGGAGTTGTCAAC 74  
 QY 71 ArgTyrSerIleArgIleGlyTyrLeuLeuIlePheArgTyrGluGlyAsnSerAlaPhe 90  
 DB 75 CGTTTCTCCATTCGGATTGGT-----TTCCAGATACAAA-----GTT 110  
 QY 91 SerValTyrIlePheAsnLeu---Ser-HisSerGluIleAsnTyrHisSerThrGly-- 108  
 DB 111 ACAGTCTACATTTTCAATTTATTCCTCCACACCTGAGATCAACACCATCTAGTAGTGA 170  
 QY 109 -----LeuMetAspSerAlaHisAsnHisPhe---LysArgAlaArgLeuPheGln 124  
 DB 171 AGCTCTTATGCAATGGATTCCGCACAGAATCAGTTCACAAACGTCGCTGATTTTGA 230  
 QY 124 uAspLeuGluAspGluAspAlaGluValIlePheProSerSerValTyrProSerProle 144  
 DB 231 AGATCTCTGAACCAAGATGCTAAGGTCAATTATTCATCGAAC----- 273

QY 144 uProGluSerThrValProAlaAsnLysGlyTyrAlaSerSerAla---IleGlnThrLe 163  
 DB 274 -CCTGATCTACTGAACCAAGTGAATAAAGGTTATGGGGTTTACAGACATCCAAAGCTT 332  
 QY 163 uPheThrGlyProValLysAlaGluGluProThrProThrProLysIleProLysLysAr 183  
 DB 333 TTTCAAA---GAATCTAAAGCTGAAGAA-----ACGCCCAAGGTACTTTAAGAAGAG 380  
 QY 183 gGlyValGlyLysLysAsnAlaAspProGluGluIleAsnSerSerAlaProArgAspAs 203  
 DB 381 AGGAGAGAGAGAGAGAGAGTCTTAATCCCGAGGAGTAATCTTCAACTCCCGGTGGAGA 440  
 QY 203 pAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrValTh 223  
 DB 441 TGACTCAGAGAACCGCTCAAAAGTTCTACGAGAGTGCTTCTGTAGAAAGAACTGTAAC 500  
 QY 223 rAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluProThrAsnProPh 243  
 DB 501 TGCAGAGAGAGAGAGAGAGCGCTCAATGCGAGCCAAACATTCGAAACCAATCCTTA 560  
 QY 243 ePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMetTyrLeuProSe 263  
 DB 561 CTTTAGAGTTGTTCTGCGACCATCATATCATATACAGAGGTGTCATCATGTACTTGCATC 620  
 QY 263 rGlyPheAlaGluLysTyrLeuSerGlyIleSerGlyPheIleLysVal 279  
 DB 621 TGGGTTTGTGAGAAATACCTAAGTGGGATATCTGNTTCATCAAGCTC 669  
 RESULT 14  
 LOCUS CD823066 645 bp mRNA linear EST 10-JUL-2003  
 DEFINITION BN25.047120F020109 BN25 Brassica napus cDNA clone BN25047120, mRNA sequence.  
 ACCESSION CD823066  
 VERSION CD823066.1 GI:32505006  
 KEYWORDS EST.  
 SOURCE Brassica napus (rape)  
 ORGANISM Brassica napus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 REFERENCE 1 (bases 1 to 645)  
 AUTHORS Genoplatte.  
 TITLE Genoplatte, a major partnership french program in plant genomics  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Genoplatte  
 Genoplatte  
 93, rue Henri Rochefort 91025 EVRY CEDEX France  
 Tel: 33 1 69 47 54 00  
 Fax: 33 1 69 47 54 10  
 This sequence has been generated in the framework of the french plant genomics programme 'Genoplatte' (<http://www.genoplatte.com>) and <http://genoplatte-info.infobiogen.fr>.  
 FEATURES  
 source  
 1. 645  
 /organism="Brassica napus"  
 /mol\_type="mRNA"  
 /cultivar="Jet Neuf"  
 /db\_xref="taxon:3708"  
 /clone="BN25047120"  
 /tissue\_type="seed"  
 /clone\_lib="BN25"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5 66e-74 Length: 645  
 Score: 786.00 Matches: 149  
 Percent Similarity: 98.70% Conservative: 3  
 Best Local Similarity: 96.75% Mismatches: 2  
 Query Match: 43.76% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-088-187a-11 (1-341) x CD823066 (1-645)

185 ArgLysLysLysAlaAspProGluGluLeuLeuAsnSerSerAlaProArgAspAsp 204  
 Db 3 AGGAGAGAGAGAACCTGTTCTGAGGAGTAAACTCATCTGCTCCGAGGGATGATGAC 62

205 ProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLysArgThrValThrAla 224  
 Db 63 CCGAGAGCGCGTTCAAGATTCTACGAGAGTCTTCTGCGAGAAAGAGACGGTTACTGCA 122

225 GluGluArgGluArgAlaAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 244  
 Db 123 GAGGAAAGAGAGAGGGCCATTATGAGCCAAACCTTCGAGCCAAACCTTTCTTC 182

245 ArgValValLeuArgProSerTyrLeuTyrArgLysLysLysLysLysLysLysLysLys 264  
 Db 183 AGAGTTGTTCTTCGACCATCTATCTATATACAGAGGCTGCATCATGTATCTGCTTCTGCT 242

265 PheAlaGluLysTyrLeuSerGlyLysLysLysLysLysLysLysLysLysLysLys 284  
 Db 243 TTTCGCTGAGAGTACTTAAAGTGGGATCTCGGGATTCATCAAGGTCCAGCTCGGGGAGAAA 302

285 GlnTyrProValArgCysLeuTyrLysAlaGlyArgAlaLysPheSerGlnGlyTyrTyr 304  
 Db 303 CAGTGGCCGGTGAGATGCTTTACAAAGCAGGAGGCCAGGTTCAGCCAGGGTGATC 362

305 GluPheThrLeuGluAsnLeuLeuGlyGluGlyAspValCysValPheGluLeuLeuArg 324  
 Db 363 GAGTTCACTCCGAGAGAACCACTAGGAGAGGTGACGTCTGCTGTTCGAGTCTCTCAGA 422

325 ThrArgAspPheValLeuLysValThrAlaPheArgValAsn 338  
 Db 423 ACCAGAGACTTCGTCTGAAAGTAAACGGCCTATCGGGTCAAC 464

## RESULT 15

CO094348

LOCUS

GR\_Eal6G14.r GR\_Ea Gossypium raimondii cDNA clone GR\_Eal6G14 3',  
 mRNA sequence.

ACCESSION

CO094348

VERSION

CO094348.1

KEYWORDS

EST.

SOURCE

Gossypium raimondii

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;  
 rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
 1 (bases 1 to 834)

REFERENCE

AUTHORS

Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,  
 Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and  
 Wing,R.A.

TITLE

Global assembly of Cotton ESTs

JOURNAL

Unpublished (2004)

COMMENT

Contact: Rod A. Wing

Arizona Genomics Institute

The University of Arizona

Forbes Building Room 303, Tucson, AZ, 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: http://genome.arizona.edu

Plate: 16 row: G column: 14.

Location/Qualifiers

1. .834

/organism="Gossypium raimondii"

/mol\_type="mRNA"

/db\_xref="taxon:29730"

/clone="GR\_Eal6G14"

/tissue\_type="whole seedlings"

/dev\_stage="first true leaves"

/lab\_host="DH10B"

/clone\_lib="GR\_Ea"

/note="Vector: pCW.SPORT-6.1; Site 1: NotI; Site 2:  
 EcoRV; library made by Invitrogen with RNA supplied by  
 Wendle lab. Directional cloned into NotI-EV. Colonies

## FEATURES

source

plated/picked by AGI. More glycerol clones held in -80."

## ORIGIN

## Alignment Scores:

Pred. No.: 2,98e-72 Length: 834  
 Score: 771.50 Matches: 171  
 Percent Similarity: 66.89% Conservativeness: 29  
 Best Local Similarity: 57.19% Mismatches: 44  
 Query Match: 42.96% Indels: 56  
 DB: 7 Gaps: 7

US-10-088-187A-11 (1-341) x CO094348 (1-834)

QY 67 GluPheValAspArgTyrSerIleArgIleGlyTyrLeuLeuIlePheArgTyrGluGly 86  
 Db 4 GAATTCCTTTCGCGTACTATATATCGCCTTTA-AAGTTTTCATCTTTCGATACGAGGG 62

QY 87 AsnSerAlaPheSerValTyrIlePheAsnLeuSerHisSerGluLeuAsnTyrHisSer 106  
 Db 63 AATTCGGCCTTTAGTGTAGTATCTTAATTTGTATATCTTCAAAATAAACTATCATGCTCT 122

QY 107 ThrGlyLeuMetAspSerAlaHisAsnHisPheLysArgAlaArgLeuPheGluAspLeu 126  
 Db 123 AATGCTCTTCTTGGTAGTCAATACATCATGGGAAA---TCATATCCGTTTCGATGAATC 179

QY 127 GluAspGluAspAlaGluValIlePheProSerSerValTyrProSerProLeuProGlu 146  
 Db 180 GAAGACGATGATGC-----194

QY 147 SerThrValProAlaAsnLysGlyTyrAlaSerSerAlaIleGlnThrLeuPheThrGly 166  
 Db 195 -----ATGCTTCAGCTATGCAGCATTTGTTTCGGGGA 227

QY 166 -----166

Db 228 TCCAAACTTAACCATTCGTTAAATTTGAGCGGTGAAGTTAACTTAAATGCAGCAAAAGC 287

QY 167 -----ProValLysAlaGluProThr-----ProThrProLysIle 179  
 Db 288 GCGAATAATCAACCTATTTCGAGTGAATTCGCTACTTCAGGTTTCAGAGAGCCACCA--- 344

QY 180 ProLysLysArgGlyArgLysLysAsnAlaAspProGluGluIleAsnSerSerAla 199  
 Db 345 CCAAGAGAACCCGGGAGGAGAAAGCAGAGTTTCGACCCCTAATGATGATGATGTTCTGTT 404

QY 200 ProArgAspAspAspProGluAsnArgSerLysPheTyrGluSerAlaSerAlaArgLys 219  
 Db 405 GGACATGAGACGATTCGGAAATCGGTATAGGTTTTCGAAAGCGCTTCAGCTCGAAAG 464

QY 220 ArgThrValThrAlaGluGluArgGluArgAlaIleAsnAlaAlaLysThrPheGluPro 239  
 Db 465 AGAACCTGACAGCCGAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 524

QY 240 ThrAsnProPhePheArgValValLeuArgProSerTyrLeuTyrArgGlyCysIleMet 259  
 Db 525 AGTAACCTTTCTGACAGAGTCGTCTTCGACCATCGTATCTATATACCGGGGATGATCATG 584

QY 260 TyrLeuProSerGlyPheAlaGluLysTyrLeuSerGlyLysPheIleLysVal 279  
 Db 585 TACCTTACCATCTTCTTTCGCCGAG---CATCTAAGTGGGGTTTCGGGGTTCTAATAACTC 641

QY 280 GlnLeuAlaGlu---LysGlnTyrProValArgCysLeuTyrLysAlaGlyArgAlaLys 298  
 Db 642 CAATCTCCGAGGGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 701

QY 299 PheSerGlnGlyTyrTyrGluPheThrLeuGluAsnAsnLeuGlyGluGlyAspValCys 318  
 Db 702 TTTAGCCAAAGATGATGATGAATTCACAGTGGAGAACAAATTTAGGGGAGGAGGAGATGTTGT 761

QY 319 ValPheGluLeuLeuArgThrArgAspPheValLeuLysValThrAlaPheArgVal 337  
 Db 762 GTCTTCGAGCTGCTTAGATCGAGGGAGGAGTCTGCTCTAAAGTGACTGTATTTCGCTGTA 818

Search completed: December 30, 2004, 09:49:31  
Job time : 6758 secs

---

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1027	100.0	1495	6	AX101036	Sequence
2	1027	100.0	1538	8	AF289052	Arabidops
3	1026	99.9	1026	6	CQ805406	Sequence
4	1025.4	99.8	1495	6	AX101038	Sequence
5	1015	98.8	1494	6	AX101039	Sequence
6	673	65.5	990	8	AY356368	Brassica
7	671.4	65.4	990	8	AY517929	Brassica
8	653.4	63.6	5000	8	AF289051	Arabidops
9	653.4	63.6	79186	8	AF000735	Arabidops
10	653	63.6	5000	6	AX101027	Sequence
11	442.6	43.1	1146	8	AY042868	Arabidops
12	397.6	38.7	681	6	AX050578	Sequence
13	270.4	26.3	125021	8	AC007504	Arabidops
14	140.4	13.7	78589	8	AP006145	Lotus cor
15	132.6	12.9	243777	2	AX137825	Medicago
16	101	9.8	2000	6	AX508278	Sequence
17	89	8.7	48368	2	AC140105	Medicago
18	83.4	8.1	114098	8	AC144459	Medicago
19	81.2	7.9	108028	8	AC135233	Medicago

Db 269 ATGCCACGCCCTTTCTTCATTAAGTTGATTTTCTCATCCTATCCAAAGAAAACGTCTG 328  
Qy 61 AGGGTCCAGATAAGTTTGTGAGTAAATCAAGGATGAGCTTTCGGTTGCTGTTGCATC 120  
Db 329 AGGGTCCAGATAAGTTTGTGAGTAAATCAAGGATGAGCTTTCGGTTGCTGTTGCATC 388  
Qy 121 ACAGTACCTGATGGTTCATGTTTGGCGGTGTAGACTAAGGAAAGCTGACAAACAAAATTTGG 180  
Db 389 ACAGTACCTGATGGTTCATGTTTGGCGGTGTAGACTAAGGAAAGCTGACAAACAAAATTTGG 448  
Qy 181 TTTCAGATGTTGGCAAGAGTTTGTGACCGTTACTCATTCGCAATTTGTTATCTTTTG 240  
Db 449 TTTCAGATGTTGGCAAGAGTTTGTGACCGTTACTCATTCGCAATTTGTTATCTTTTG 508  
Qy 241 ATTTTATAGATGAGGAACTCTCCCTTCAGGCTACATTTTCAATTTATCCACTCT 300  
Db 509 ATTTTATAGATGAGGAACTCTCCCTTCAGGCTACATTTTCAATTTATCCACTCT 568  
Qy 301 GAGATCAATTACATTTCCACCGGTCTCATGATTCGGCTCAACACCACTTCAACGCGCC 360  
Db 569 GAGATCAATTACATTTCCACCGGTCTCATGATTCGGCTCAACACCACTTCAACGCGCC 628  
Qy 361 CGTTTGTGAGAGCTTGAAGATCAAGATGCGGAGGTCACTTTCTCTTCTCTGTGATC 420  
Db 629 CGTTTGTGAGAGCTTGAAGATCAAGATGCGGAGGTCACTTTCTCTTCTCTGTGATC 688  
Qy 421 CCATCACCACCTTCTGAGTCTACAGTCCAGCCCAACAAAGGTATGCTAGTTCAAGCATC 480  
Db 689 CCATCACCACCTTCTGAGTCTACAGTCCAGCCCAACAAAGGTATGCTAGTTCAAGCATC 748  
Qy 481 CAAACCTTGTTCACCTGGACAGTAAAGCTGAAGAGCCAAAGCCCAACCCCAAAATACCT 540  
Db 749 CAAACCTTGTTCACCTGGACAGTAAAGCTGAAGAGCCAAAGCCCAACCCCAAAATACCT 808  
Qy 541 AAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
Db 809 AAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 868  
Qy 601 CGAGATGATGATCCAGAGAACCGTTCAAGATTTACGAGAGTCTTCGCGAGGAGGAGGAG 660  
Db 869 CGAGATGATGATCCAGAGAACCGTTCAAGATTTACGAGAGTCTTCGCGAGGAGGAGGAG 928  
Qy 661 ACCGTGACTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
Db 929 ACCGTGACTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 988  
Qy 721 AACCCCTTTCTCAGAGTGGTTCTGCGAACCATCTATCTATACAGAGGTTGCAATCATGAT 780  
Db 989 AACCCCTTTCTCAGAGTGGTTCTGCGAACCATCTATCTATACAGAGGTTGCAATCATGAT 1048  
Qy 781 CTTCTCTTCTGGGTTTCTGAGAGTACCTAAGTGGGATCTCCGGTTTCATCAAGTCCAG 840  
Db 1049 CTTCTCTTCTGGGTTTCTGAGAGTACCTAAGTGGGATCTCCGGTTTCATCAAGTCCAG 1108  
Qy 841 CTTGCGGAGAAACAATGGCTCTGTCATCTCTCAAAAGCCGAGAGCCAAATTCAGT 900  
Db 1109 CTTGCGGAGAAACAATGGCTCTGTCATCTCTCAAAAGCCGAGAGCCAAATTCAGT 1168  
Qy 901 CAAGGATGATCAAGAAATTCATCTAGAGAAACAATTAGGAGAGGAGAGCGTCTGTGTGTT 960  
Db 1169 CAAGGATGATCAAGAAATTCATCTAGAGAAACAATTAGGAGAGGAGAGCGTCTGTGTGTT 1228  
Qy 961 GAGTCTCAGAACCAAGATTTGTTTGAAGTGAAGCCCTTCGAGTCAACAGTAC 1020  
Db 1229 GAGTCTCAGAACCAAGATTTGTTTGAAGTGAAGCCCTTCGAGTCAACAGTAC 1288  
Qy 1021 GTCTGAA 1027  
Db 1289 GTCTGAA 1295

AF289052 1538 bp mRNA linear PLN 05-NOV-2003  
LOCUS Arabidopsis thaliana reduced vernalization response 1 (VRN1) mRNA,  
DEFINITION VRN1-Ler allele, complete cds.  
ACCESSION AF289052  
VERSION AF289052.1 GI:21734795  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
REFERENCE 1 (bases 1 to 1538)  
AUTHORS Levy, Y., Messing, S., Mylne, J.S., Gendall, A.R. and Dean, C.  
TITLE Multiple roles of Arabidopsis VRN1 in vernalization and flowering  
time control  
JOURNAL Science 297 (5579), 243-246 (2002)  
MEDLINE 22111275  
PUBMED 12114624  
REFERENCE 2 (bases 1 to 1538)  
AUTHORS Levy, Y., Gendall, A.R. and Dean, C.  
TITLE VRN1, a gene required for response to vernalization  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 1538)  
AUTHORS Levy, Y. and Dean, C.  
TITLE Direct Submission  
JOURNAL Submitted (23-JUL-2000) Molecular Genetics, John Innes Centre,  
Colney Lane, Norwich, Norfolk NR4 7UH, UK  
FEATURES  
source  
1..1538  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/map="between mi339 and m560B2"  
/ecotype="Landsberg erecta"  
1..1538  
/gene="VRN1"  
/allele="Ler"  
312..1337  
/gene="VRN1"  
/allele="Ler"  
/note="contains two VPI/ABI3-like B3 domains"  
/product="reduced vernalization response 1"  
/protein\_id="AAW76973.1"  
/db\_xref="GI:21734795"  
/translation="MPRPFFPKLIFSSITQKRLRVPDKFVKFDELVAVALTVPD  
GHVVRGLRKADNKIWFQDQWQEFVDYRISIRIGYLLIFRYEGNSAFSVYIENLHSEI  
NYHSTGLMDSAHNHFKARLPDELDEDAEVIFFPSVYFPLPSTVPANKYASSAI  
QTLFTGPVKAEEPTPKI PKKGRKKKNADPEEINSAPRDDDPENRKFVGSASAR  
KRTVTAERERAINAAKTFPTNPFVRLPSYLRGCIIVLPSPGAEKYLSTGSGF  
IKVQLAEQWPRCLYKAGRAKFSQGWYEFLENNLGECDVCVFELLRTDFVLKVTA  
FRVNEYV"  
ORIGIN  
Query Match 100.0%; Score 1027; DB 8; Length 1538;  
Best Local Similarity 100.0%; Pred. No. 2.3e-287;  
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGCCACGCCCTTTCTCCATAAGTTGATTTTCTCATCCTATCCAAAGAAAACGTCTG 60  
Db 312 ATGCCACGCCCTTTCTCCATAAGTTGATTTTCTCATCCTATCCAAAGAAAACGTCTG 371  
Qy 61 AGGGTCCAGATAAGTTTGTGAGTAAATCAAGGATGAGCTTTCGGTTGCTGTTGCATC 120  
Db 372 AGGGTCCAGATAAGTTTGTGAGTAAATCAAGGATGAGCTTTCGGTTGCTGTTGCATC 431  
Qy 121 ACAGTACCTGATGGTTCATGTTTGGCGGTGTAGACTAAGGAAAGCTGACAAACAAAATTTGG 180  
Db 432 ACAGTACCTGATGGTTCATGTTTGGCGGTGTAGACTAAGGAAAGCTGACAAACAAAATTTGG 491  
Qy 181 TTTCAGATGTTGGCAAGAGTTTGTGACCGTTACTCATTCGCAATTTGTTATCTTTTG 240

492 TTTCAAGATGGTGGCAAGATTTGTTGACCGTACTCCATTCGATGTTATCTTTTG 551  
241 ATTTTATAGATATGAAGAAATCTGCTTTCAGGCTCTACATTTTCAATTTATCCCACTCT 300  
552 ATTTTATAGATATGAAGAAATCTGCTTTCAGGCTCTACATTTTCAATTTATCCCACTCT 611  
301 GAGATCAATATACATTTCCACCGGCTCTCATGATTTCCGCTTCAACACCATTTCAACCGGCC 360  
612 GAGATCAATATACATTTCCACCGGCTCTCATGATTTCCGCTTCAACACCATTTCAACCGGCC 671  
361 CGTTTGTGTTGAGACCTTGAAGATGAAGATGCGGAGTTCATCTTTCTTTCTGTGTAC 420  
672 CGTTTGTGTTGAGACCTTGAAGATGAAGATGCGGAGTTCATCTTTCTTTCTGTGTAC 731  
421 CCATCACCACTTCTGAGTCTACAGTACCAGCCCAACAAAGGATGCTAGTTTCAGCCATC 480  
732 CCATCACCACTTCTGAGTCTACAGTACCAGCCCAACAAAGGATGCTAGTTTCAGCCATC 791  
481 CAAACCTTGTTCATCTGACCAAGTTAAAGCTGAAGAGCCCAACGCCCAACCAAAATACCT 540  
792 CAAACCTTGTTCATCTGACCAAGTTAAAGCTGAAGAGCCCAACGCCCAACCAAAATACCT 851  
541 AAAAGAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
852 AAAAGAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 911  
601 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGAGTCTTCTGCGAGAGAGAGA 660  
912 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGAGTCTTCTGCGAGAGAGAGA 971  
661 ACCGTGACTGCGAG 720  
972 ACCGTGACTGCGAG 1031  
721 AACCTTCTTTCAGAGTGTCTTCTGAGACCATCTATCTATACAGAGGTTGATCATGTAT 780  
1032 AACCTTCTTTCAGAGTGTCTTCTGAGACCATCTATCTATACAGAGGTTGATCATGTAT 1091  
781 CTTCTCTTCTGAGTGTCTGAGAGTACCTTAAGTGGATCTCCGGTTCATCAAGTCCAG 840  
1092 CTTCTCTTCTGAGTGTCTGAGAGTACCTTAAGTGGATCTCCGGTTCATCAAGTCCAG 1151  
841 CTTGCGGAG 900  
1152 CTTGCGGAG 1211  
901 CAGAGATGATCAATTTCTATAG 960  
1212 CAGAGATGATCAATTTCTATAG 1271  
961 GAGTGTCTCAG 1020  
1272 GAGTGTCTCAG 1331  
1021 GTCTGAA 1027  
1332 GTCTGAA 1338

RESULT 3  
LOCUS CQ805406  
DEFINITION Sequence 1817 from Patent WO2004035798.  
ACCESSION CQ805406  
VERSION CQ805406.1 GI:47111337  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1

AUTHORS Inze,D., de Veylder,L. and Vlieghe,K.  
TITLE Identification of novel e2f target genes and use thereof  
JOURNAL Patent: WO 2004035798-A 1817 29-APR-2004;  
CropDesign N.V. (BE)  
FEATURES  
source Location/Qualifiers  
1..1026  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"  
ORIGIN  
Query Match 99.9%; Score 1026; DB 6; Length 1026;  
Best Local Similarity 100.0%; Pred. No. 4.3e-287;  
Matches 1026; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGCCACGCCCTTCTTCATTAAGTTGATTTTCTCATCCACTATCCAGAAAAACGCTCG 60  
DB 1 ATGCCACGCCCTTCTTCATTAAGTTGATTTTCTCATCCACTATCCAGAAAAACGCTCG 60  
QY 61 AGGGTCCAGATAAGTTTGTGAGTAAATTTCAAGATGAGCTTTCGGTTGCTGTGCATCTC 120  
DB 61 AGGGTCCAGATAAGTTTGTGAGTAAATTTCAAGATGAGCTTTCGGTTGCTGTGCATCTC 120  
QY 121 ACAGTACCTGATGTTTGGCGTGTAGACTAAGAAAGCTGACAAACAAATTTGG 180  
DB 121 ACAGTACCTGATGTTTGGCGTGTAGACTAAGAAAGCTGACAAACAAATTTGG 180  
QY 181 TTTCAAGATGTTTGGCAAGAGTTTCTGACCGTTACTCCATTCGATGTTTATCTTTTGG 240  
DB 181 TTTCAAGATGTTTGGCAAGAGTTTCTGACCGTTACTCCATTCGATGTTTATCTTTTGG 240  
QY 241 ATTTTATAGATATGAAGAAATCTCTGCCCTTACAGCTCTACATTTTCAATTTATCCCACTCT 300  
DB 241 ATTTTATAGATATGAAGAAATCTCTGCCCTTACAGCTCTACATTTTCAATTTATCCCACTCT 300  
QY 301 GAGATCAATATACCTTCCACCGGCTCTCATGATTCGCTCAACACATTTCAAAACGCGCC 360  
DB 301 GAGATCAATATACCTTCCACCGGCTCTCATGATTCGCTCAACACATTTCAAAACGCGCC 360  
QY 361 CGTTTGTGTTGAGACCTTGAAGATGAAGATGCGGAGTTCATCTTCTTCTTCTGTGTAC 420  
DB 361 CGTTTGTGTTGAGACCTTGAAGATGAAGATGCGGAGTTCATCTTCTTCTTCTGTGTAC 420  
QY 421 CCATCACCACTTCTGAGTCTACAGTACCGAGCAACAAAGGATGCTAGTTTCAGCCATC 480  
DB 421 CCATCACCACTTCTGAGTCTACAGTACCGAGCAACAAAGGATGCTAGTTTCAGCCATC 480  
QY 481 CAAACCTTGTTCATCTGAGCCAGTTAAAGCTGAAGAGCCAAACGCCCAACCAAAATACCT 540  
DB 481 CAAACCTTGTTCATCTGAGCCAGTTAAAGCTGAAGAGCCAAACGCCCAACCAAAATACCT 540  
QY 541 AAAAGAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
DB 541 AAAAGAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
QY 601 CGAGATGATGATCCAGAGAACCGTTTCAAAGTTCTACGAGAGTCTTCTGCGAGAGAGAGA 660  
DB 601 CGAGATGATGATCCAGAGAACCGTTTCAAAGTTCTACGAGAGTCTTCTGCGAGAGAGAGA 660  
QY 661 ACCGTGATGCGAG 720  
DB 661 ACCGTGATGCGAG 720  
QY 721 AACCCCTTCTCAGAGTGTCTGCGACCATCTATCTATACAGAGGTTGATCATGTAT 780  
DB 721 AACCCCTTCTCAGAGTGTCTGCGACCATCTATCTATACAGAGGTTGATCATGTAT 780  
QY 781 CTTCTCTTCTGAGTGTCTGAGAGTACCTTAAGTGGATCTCCGGTTCATCAAGTCCAG 840  
DB 781 CTTCTCTTCTGAGTGTCTGAGAGTACCTTAAGTGGATCTCCGGTTCATCAAGTCCAG 840  
QY 841 CTTGCGGAG 900  
DB 841 CTTGCGGAG 900

```
Db      841  CTTGCGGAGAAACAATGGCGTGTTCGATGCTCTCAAAAGCCGGAGAGCAAAATTCAGT 900
Qy      901  CRAAGTGTGTACAAATTCCTCTAGAGAACACTTAGGAGAGGAGACGCTCTGTGTGTTT 960
Db      901  CRAAGTGTGTACAAATTCCTCTAGAGAACACTTAGGAGAGGAGACGCTCTGTGTGTTT 960
Qy      961  GAGCTGCTCAGAACACAGAGATTTGTTTTGAAAGTGACAGCCCTTTTCAGTCAACGAGTAC 1020
Db      961  GAGCTGCTCAGAACACAGAGATTTGTTTTGAAAGTGACAGCCCTTTTCAGTCAACGAGTAC 1020
Qy      1021  GTCTGA 1026
Db      1021  GTCTGA 1026

RESULT 4
AX101038
LOCUS      AX101038      1495 bp      DNA      linear      PAT 10-APR-2001
DEFINITION Sequence 12 from Patent WO0121822.
ACCESSION AX101038
VERSION    AX101038.1 GI:13619894
KEYWORDS   .
SOURCE      synthetic construct
            synthetic construct
            artificial sequences.
ORIGIN      1
REFERENCE   1
AUTHORS     Dean,C. and Levy,Y.Y.
TITLE       Methods and means for modification of plant flowering
            Characteristics
JOURNAL     Patent: WO 0121822-A 12 29-MAR-2001;
            Plant Bioscience Limited (GB)
FEATURES    Location/Qualifiers
            1..1495
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="vrnl-1 mutation"

ORIGIN
Query Match      99.8%; Score 1025.4; DB 6; Length 1495;
Best Local Similarity 99.9%; Pred. No. 6.7e-287;
Matches 1026; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  ATGCCACGCCCTTCTTCCATAAGTTGATTTCTCATCCACTATCCAGAAAAACGTCG 60
Db      269  ATGCCACGCCCTTCTTCCATAAGTTGATTTCTCATCCACTATCCAGAAAAACGTCG 328
Qy      61  AGGTCCTCCAGATAAGTTGTGAGTAAATTCAGGATGAGCTTCGGTTGCTGTGCACATC 120
Db      329  AGGTCCTCCAGATAAGTTGTGAGTAAATTCAGGATGAGCTTCGGTTGCTGTGCACATC 388
Qy      121  ACAGTACCTGATGGTCATCTTTGGCGGTGAGGACTAAGGAAAGCTGACAAACAAATTTGG 180
Db      389  ACAGTACCTGATGGTCATCTTTGGCGGTGAGGACTAAGGAAAGCTGACAAACAAATTTGG 448
Qy      181  TTTCAAGATGGTTGGCAAGATTTGTTGACCGTTACTCCATTGCGATGGTTATCTTTTG 240
Db      449  TTTCAAGATGGTTGGCAAGATTTGTTGACCGTTACTCCATTGCGATGGTTATCTTTTG 508
Qy      241  ATTTTATAGATGAAGAAACTCTGCTTCAGCGTCAATTTTCAATTTATCCACTCT 300
Db      509  ATTTTATAGATGAAGAAACTCTGCTTCAGCGTCAATTTTCAATTTATCCACTCT 568
Qy      301  GAGATCAATTTACCATTTCCACCGGCTCATGGATTCGGCTCAACACCTTCAACGCGCC 360
Db      569  GAGATCAATTTACCATTTCCACCGGCTCATGGATTCGGCTCAACACCTTCAACGCGCC 628
Qy      361  CGTTTGTGTTGAAGACCTTGAAGATGAAGTGGCGAGGTCAATCTTTCTCTCTGTGTAC 420
Db      629  CGTTTGTGTTGAAGACCTTGAAGATGAAGTGGCGAGGTCAATCTTTCTCTCTGTGTAC 688
Qy      421  CATACACCATCTCTGAGTCTACAGTACGAGCCCAAAAGGGTATGCTAGTTTCAGGCATC 480
```

```
Db      689  CCATCACCACCTTCTCTGAGTCTACAGTACGAGCAACAAAGGATGTAGTTTCAGCCATC 748
Qy      481  CAAACCTTTGTTCACTGGACCAAGTTAAAGCTGAAGAGCCAAACGCCAACCCCAAAATACCT 540
Db      749  CAAACCTTTGTTCACTGGACCAAGTTAAAGCTGAAGAGCCAAACGCCAACCCCAAAATACCT 808
Qy      541  AAAAAGAGAGGAGGAGGAAGAAATAATGCTGATCTCTGAGGAAATAAACTCATCAGCTCG 600
Db      809  AAAAAGAGAGGAGGAGGAAGAAATAATGCTGATCTCTGAGGAAATAAACTCATCAGCTCG 868
Qy      601  CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGAGTGTCTTCGCGAGAAAGAGA 660
Db      869  CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGAGTGTCTTCGCGAGAAAGAGA 928
Qy      661  ACCGTGACTGAGAGAAAGAGAGAGAGCCCATCAATGACGACCAAAACCTTCGAACCAACA 720
Db      929  ACCGTGACTGAGAGAAAGAGAGAGAGAGCCCATCAATGACGACCAAAACCTTCGAACCAACA 988
Qy      721  AACCCCTTTCTTCAGAGTGGTTCTGCGACCATCTCTATCTATACAGAGGTTGCATCATGTAT 780
Db      989  AACCCCTTTCTTCAGAGTGGTTCTGCGACCATCTCTATCTATACAGAGGTTGCATCATGTAT 1048
Qy      781  CTTCTCTCTGGGTTTGTGAGAGTACCTAAAGTGGATCTCCGGTTTCATCAAAAGTCCAG 840
Db      1049  CTTCTCTCTGGGTTTGTGAGAGTACCTAAAGTGGATCTCCGGTTTCATCAAAAGTCCAG 1108
Qy      841  CTTGCGGAGAAACAATGGCGTGTTCGATGCTCTTACAAGCGGGGAGAGCCAAATTCAGT 900
Db      1109  CTTGCGGAGAAACAATGGCGTGTTCGATGCTCTTACAAGCGGGGAGAGCCAAATTCAGT 1168
Qy      901  CAAGATGCTACGAATTCACCTTAGAGAAACAATTTAGGAGAGGAGAGAGCTCTGTGTGTTT 960
Db      1169  CAAGATGCTACGAATTCACCTTAGAGAAACAATTTAGGAGAGGAGAGAGCTCTGTGTGTTT 1228
Qy      961  GAGCTGCTCAGAACACAGAGATTTGTTTTGAAAGTGACAGCTTCGAGTCAACGAGTAC 1020
Db      1229  GAGCTGCTCAGAACACAGAGATTTGTTTTGAAAGTGACAGCTTCGAGTCAACGAGTAC 1288
Qy      1021  GTCTGA 1027
Db      1289  GTCTGA 1295

RESULT 5
AX101039
LOCUS      AX101039      1494 bp      DNA      linear      PAT 10-APR-2001
DEFINITION Sequence 13 from Patent WO0121822.
ACCESSION AX101039
VERSION    AX101039.1 GI:13619895
KEYWORDS   .
SOURCE      synthetic construct
            synthetic construct
            artificial sequences.
ORIGIN      1
REFERENCE   1
AUTHORS     Dean,C. and Levy,Y.Y.
TITLE       Methods and means for modification of plant flowering
            Characteristics
JOURNAL     Patent: WO 0121822-A 13 29-MAR-2001;
            Plant Bioscience Limited (GB)
FEATURES    Location/Qualifiers
            1..1494
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="vrnl-2 mutation"

ORIGIN
Query Match      98.8%; Score 1015; DB 6; Length 1494;
Best Local Similarity 99.9%; Pred. No. 7.1e-284;
Matches 1026; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy      1  ATGCCACGCCCTTCTTCCATAAGTTGATTTCTCATCCACTATCCAGAAAAACGTCG 60
```

Db 269 ATGCCAGGCCCTTTCTTCCATAAGTTGATTTTCTCATCCATATCCAGAAAAACGCTCTG 328  
QY 61 AGGGTCCAGATAAGTTTGTGAGTAAATTCAGAGATGAGCTTTCCGTTGCTGTGGCACTC 120  
Db 329 AGGGTCCAGATAAGTTTGTGAGTAAATTCAGAGATGAGCTTTCCGTTGCTGTGGCACTC 388  
QY 121 ACAGTACCTGATGTCATGTTTCCGCTGTAGGACTAAGGAAGCTGACACAAAAATTTGG 180  
Db 389 ACAGTACCTGATGTCATGTTTCCGCTGTAGGACTAAGGAAGCTGACACAAAAATTTGG 448  
QY 181 TTTCAAGATGTTGGCAAGATTTGTGACCGTTACTCCATTCGCATTCGTTATCTTTTGG 240  
Db 449 TTTCAAGATGTTGGCAAGATTTGTGACCGTTACTCCATTCGCATTCGTTATCTTTTGG 508  
QY 241 ATTTTATAGATATCAAGAAACTCTGCTTCCAGGCTTACATTTTCAATTTATCCCACTCT 300  
Db 509 ATTTTATAGATATCAAGAAACTCTGCTTCCAGGCTTACATTTTCAATTTATCCCACTCT 568  
QY 301 GAGATCAATTAACATTCACCGCTCTCATGATTCGCTTCAACCACTTCAAAACGCGCC 360  
Db 569 GAGATCAATTAACATTCACCGCTCTCATGATTCGCTTCAACCACTTCAAAACGCGCC 628  
QY 361 CGTTTCTTTGAAGACCTTGAAGATGAAGATGCGAGTGCATCTTTCTTCTCTGTTAC 420  
Db 629 CGTTTCTTTGAAGACCTTGAAGATGAAGATGCGAGTGCATCTTTCTTCTCTGTTAC 688  
QY 421 CCATCACCACTTCTCTGAGTCTACAGTACCAGCCAAACAAAGGGTATGCTAGTTCCAGCCATC 480  
Db 689 CCATCACCACTTCTCTGAGTCTACAGTACCAGCCAAACAAAGGGTATGCTAGTTCCAGCCATC 748  
QY 481 CAAACCTTGTCTCTGAGCCAGTAAAGTGAAGGCCAAACGCGCAACCCCAAAAAATACCT 540  
Db 749 CAAACCTTGTCTCTGAGCCAGTAAAGTGAAGGCCAAACGCGCAACCCCAAAAAATACCT 808  
QY 541 AAAAGAGAGGGAGGAGAAATGCTGATCTCTGAGGAATAAATCACTCATCAGCTCCG 600  
Db 809 AAAAGAGAGGGAGGAGAAAG-AAAATGCTGATCTCTGAGGAATAAATCACTCATCAGCTCCG 867  
QY 601 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTAGCAGAGTCTTCTCGAGAAAGAGA 660  
Db 868 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTAGCAGAGTCTTCTCGAGAAAGAGA 927  
QY 661 ACCGTGACTGCGAGAAAGAGAGAGAGCCATCAATGCGAGCCAAACGTTGGAACCAACA 720  
Db 928 ACCGTGACTGCGAGAAAGAGAGAGAGCCATCAATGCGAGCCAAACGTTGGAACCAACA 987  
QY 721 AACCCTTTCTCAGAGTGGTTCTGCGACCATCTCTATACAGAGGTTGCATCATGTAT 780  
Db 988 AACCCTTTCTCAGAGTGGTTCTGCGACCATCTCTATACAGAGGTTGCATCATGTAT 1047  
QY 781 CTTCTCTTCTGGGTTTCTGAGAGTACCTAAGTGGGATCTCGGGTTCTATCAAGTCCAG 840  
Db 1048 CTTCTCTTCTGGGTTTCTGAGAGTACCTAAGTGGGATCTCGGGTTCTATCAAGTCCAG 1107  
QY 841 CTTGCGGAGAAACAAATGGCTGTTCGATGTCTCTACAAAGCCGGAGAGCCAAATTCAGT 900  
Db 1108 CTTGCGGAGAAACAAATGGCTGTTCGATGTCTCTACAAAGCCGGAGAGCCAAATTCAGT 1167  
QY 901 CAAGGATGTACGAATTCATCTAGAGAACCACTTAGGAGAGGAGAGCGTCTGTGTGTTT 960  
Db 1168 CAAGGATGTACGAATTCATCTAGAGAACCACTTAGGAGAGGAGAGCGTCTGTGTGTTT 1227  
QY 961 GAGCTCTCAGAACACAGAGATTTGTTTTGAAAGTGCAGAGCCCTTTCGAGTCAACGAGTAC 1020  
Db 1228 GAGCTCTCAGAACACAGAGATTTGTTTTGAAAGTGCAGAGCCCTTTCGAGTCAACGAGTAC 1287  
QY 1021 GTCTGAA 1027  
Db 1288 GTCTGAA 1294

RESULT 6  
AY356368

LOCUS AY356368 990 bp mRNA linear PLN 25-AUG-2003  
DEFINITION Brassica rapa cultivar Samjin reduced vernalization response 1  
mRNA, complete cds.  
ACCESSION AY356368  
VERSION AY356368.1 GI:33943516  
KEYWORDS  
SOURCE Brassica rapa  
ORGANISM Brassica rapa  
REFERENCE  
AUTHORS Kwon,S.-J., Park,B.-S., Kim,S.-Y., Choi,H.-S., Lee,M.-C.,  
Kim,J.-S., Lee,S.-I., Lim,K.-B., Kim,J.-A., Hong,K.-Y., Lee,M.-R.,  
Jin,Y.-M., Kim,D. and Kim,H.-I.  
TITLE Direct Submission  
JOURNAL Submitted (23-JUL-2003) Genomics Division, National Institute of  
Agricultural Biotechnology, Seodun-dong 225, Suwon 441-707,  
Republic of Korea  
FEATURES  
source  
1..990  
/organism="Brassica rapa"  
/mol\_type="mRNA"  
/cultivar="Samjin"  
/db\_xref="taxon:3711"  
CDS  
1..990  
/note="VRN1"  
/codon\_start=1  
/product="reduced vernalization response 1"  
/protein\_id="AAQ55453.1"  
/db\_xref="GI:33943517"  
/translation="MPPFFHKLIFSSITQEKRLRVPKFKDELSVAVALTVPD  
EHWVRLKADNNKINWQDQBFVDRYSIRIGYLLIFRYEGNSAFVCIYNLPQS  
GHNHSTGLMDSASHNNHFKRPLFEDLEDEDAETLHTTASAIQSPFTGPKPEBATP  
TQTSKVPKGRKKXNADHPVEVNSAPRDDPESRSKEYESASARKGTNABERRA  
VNAKTFEPTNPFVRLRPSLYLPSGFAEKYLSGSIKFKVQLGSKMPV  
RCLYKAGRAKFSQGWYBFTVENNLGEGDVCVFELLRTDFVLKVTAYRVNYY"

## ORIGIN

Query Match 65.5%; Score 673; DB 8; Length 990;  
Best Local Similarity 82.5%; Pred. No. 2,4e-184; Mismatches 105; Indels 78; Gaps 5;  
Matches 864; Conservative 0;

QY 1 ATGCCACGCGCTTCTTCCATAAGTTGATTTTCTCATCCATATCCAGAAAAACGCTCTG 60  
Db 1 ATGCCACGCGCTTCTTCCACAGCTGATTTTCTCATCCATATCCAGAAAAACGCTCTG 60  
QY 61 AGGTCCTCCAGATAAGTTTGTGAGTAAATTCAGAGATGAGCTTTCGTTGTTGCATCTC 120  
Db 61 AGAGTTCAGATAAGTTTGTGAGTAAATTCAGAGACGAGCTATCGTTGCGTTCATCTC 120  
QY 121 ACAGTACCTGATGTCATGTTTGGCGGTGAGGACTAAGCAAGCTG-----ACAACAA 174  
Db 121 ACAGTACCTGATGTCATGTTTGGCGGTGAGGACTAAGCAAGCTG-----ACAACAA 180  
QY 175 ATTTGGTTTCAAGATGTTTGGCAAGAGTTTGTGACCGTTACTCTCAATTCGCGATTGTTAT 234  
Db 181 ATTTGGTTTCAAGATGTTTGGCAAGAGTTTGTGACCGTTACTCTCAATTCGCGATTGTTAT 240  
QY 235 CTTTGTATTTTATAGATATGAAGAACTCTGCTTCAGCGTCTACATTTTCAATTTATCC 294  
Db 241 CTTTGTATTTTATAGATATGAAGAACTCTGCTTCAGCGTCTGCACTTTTACAACTTACCA 300  
QY 295 CACTCTGAGATCAATTAACCTTCCACCGGTCTCATGGATTCGCG-----TCACACACCAC 348  
Db 301 CAGTCCGAGATTAACCTTCCACCGGTCTCATGGACTCTGCACTACACACACACACCAC 360  
QY 349 TTTCAAAACGCGCGCTTTGTTTCAAGACCTTGAAGATGAAGATGCGAGGTCATCTTTCTCT 408  
Db 361 TTTCAAAACGCGCGCTTTGTTTCAAGACCTTGAAGATGAAGATGCGAGGTCATCTGCAACA 420  
QY 409 TCTTCTGTGTACCCACTCACCACTTCTCTGAGTCTACAGTACCAGCCAAACAAAGGTTATGCT 468



```
QY 760 TACAGAGTTGCATCATGTATCTCTCTCTGCGTTGCTGAGAGTACCTAAAGTGGATC 819
Db 724 TACAGAGTTGCATCATGTATCTCTCTCTGCGTTGCTGAGAGTACCTAAAGTGGATC 783
QY 820 TCCGGGTTTCATCAAGTCCAGCTTCGGGAGAACAAATGGCTGCTCGATGCTCTACAAA 879
Db 784 TCCGGGTTTCATCAAGTCCAGCTTCGGGAGAACAAATGGCTGCTCGATGCTCTACAAA 843
QY 880 GCGGGAGAGCCAAATTCAGTCAAGGATGCTACGAATTCACCTACAGAACCACTTAGGA 939
Db 844 GCGGGAGAGCTAAAGTTCAGTCAAGGATGCTACGAATTCACCTACAGAACCACTTAGGA 903
QY 940 GAAGGAGAGCTCTGTGTTGAGCTGCTCAGAACAGAGATTTGCTTTTGAAGTGACA 999
Db 904 GAAGGAGAGCTCTGTGTTGAGCTGCTCAGAACAGAGATTTGCTTTTGAAGTGAGC 963
QY 1000 GCCTTTTCAGTCAACAGTACGTCGA 1026
Db 964 GCCTATCGAGTCAACAGTACGTCGA 990

RESULT 8
AF289051
LOCUS
DEFINITION
Arabidopsis thaliana 5000 bp DNA linear PLN 05-NOV-2003
VRN1-Ler allele, complete cds.
ACCESSION
AF289051
VERSION
AF289051.1 GI:21734793
KEYWORDS
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 5000)
Levy,Y.Y., Mesnage,S., Wylne,J.S., Gendall,A.R. and Dean,C.
Multiple roles of Arabidopsis VRN1 in vernalization and flowering
time control
Science 297 (5579), 243-246 (2002)
JOURNAL
MEDLINE
22111275
PUBMED
22114624
REFERENCE
2 (bases 1 to 5000)
Levy,Y.Y., Gendall,A.R. and Dean,C.
VRN1, a gene required for response to vernalization
Unpublished
3 (bases 1 to 5000)
Levy,Y.Y. and Dean,C.
Direct Submission
Submitted (23-JUL-2000) Molecular Genetics, John Innes Centre,
Colney Lane, Norwich, Norfolk NR4 7UH, UK
FEATURES
Source
1..5000
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="3"
/map="between mi339 and m560B2"
/ecotype="Landsberg erecta"
<1879..>4370
/gene="VRN1"
/allele="Ler"
join(<1879..1938,3120..3567,3664..3732,3832..4031,
4122..>4370)
/gene="VRN1"
/allele="Ler"
join(1879..1938,3120..3567,3664..3732,3832..4031,
4122..4370)
/gene="VRN1"
/allele="Ler"
/note="contains two VPI/ABI3-like B3 domains"
/codon_start=1
/product="reduced vernalization response 1"
/protein_id="AAM76972.1"
```

```
/db_xref="GI:21734794"
/translation="MPRPFFHKLIFSSTIQEKLRLVPDKFVKFDELSSVAVALTPD
GHVWVGLRKADNKIWFQDMQFVDRYSIRIGYLLIFRGNSSAFSVYIFNLSHSEI
NVHSTGLMDSAHNHFKRALPEDLEDAEIVPPSSVPSPLPESTVPANKGVASSAI
OTLFTGPVKAERTPTPKIPKRGKKNADPEEINSSAPRDDDDPENRSKFVESASAR
KXIVTAEBRERAINAAKFPNPFVRLRPSYLYRGICIMYLPSPGFASKYLSGISGF
IKVQAEQWMPVRCLYKAGRAKFSQGWVEFTLENNILGEGDVCFELLRTDRDFVLVTA
FRVNEYV"
ORIGIN
Query Match 63.6%; Score 653.4; DB 8; Length 5000;
Best Local Similarity 77.2%; Pred. No. 1.5e-178;
Matches 969; Conservative 0; Mismatches 1; Indels 285; Gaps 3;
QY 58 CTGAGGGTCCCAAGATTAAGTTTGTGAGTAAATCAAGATGAGCTTTCCGGTTGCTGTGCA 117
Db 3117 CAGAGGGTCCCAAGATTAAGTTTGTGAGTAAATCAAGATGAGCTTTCCGGTTGCTGTGCA 3176
QY 118 CTCACAGTACCTGATGGTCATGTTTGGCGGTGAGGCTAAGGAAAGCTGACCAAAAATT 177
Db 3177 CTCACAGTACCTGATGGTCATGTTTGGCGGTGAGGCTAAGGAAAGCTGACCAAAAATT 3236
QY 178 TGGTTTCAAGATGGTTGGCAAGAGTTTGTGACCGTTACTCCATTCGCATTGGTTATCTT 237
Db 3237 TGGTTTCAAGATGGTTGGCAAGAGTTTGTGACCGTTACTCCATTCGCATTGGTTATCTT 3296
QY 238 TTGATTTTGTAGATATGAAGAAACTCTGCTTTCAGCGTCTACATTTTCAATTTATCCCAC 297
Db 3297 TTGATTTTGTAGATATGAAGAAACTCTGCTTTCAGCGTCTACATTTTCAATTTATCCCAC 3356
QY 298 TCTGAGATCAATTACCATTCCACCGGTCTCATGGATTCGCTCACAAACCACTTCAACGC 357
Db 3357 TCTGAGATCAATTACCATTCCACCGGTCTCATGGATTCGCTCACAAACCACTTCAACGC 3416
QY 358 GCCCGTTGTTTGAAGACCTTGAAGATGAAGTGGCAGGTCATCTTCTTCTTCTG 417
Db 3417 GCCCGTTGTTTGAAGACCTTGAAGATGAAGTGGCAGGTCATCTTCTTCTTCTG 3476
QY 418 TACCATACCACTTCTCTGAGTCTACAGTACCAAGCAAAAGGGTATGCTAGTTCAGCC 477
Db 3477 TACCATACCACTTCTCTGAGTCTACAGTACCAAGCAAAAGGGTATGCTAGTTCAGCC 3536
QY 478 ATCCAAACCTTCTCTGAGTCTACAGTACCAAGCAAAAGGGTATGCTAGTTCAGCC 506
Db 3537 ATCCAAACCTTCTCTGAGTCTACAGTACCAAGCAAAAGGGTATGCTAGTTCAGCC 3596
QY 507 ----- 506
Db 3597 ATCTATCGTGATTAAGCGTCTTATCTTTTGGGTTGATGCTGTGATATTTTCTT 3656
QY 507 -----AGCTGAAGAGCCACGCCAACCCCAAAATACCTTAAAGAGAGAGGAGGAAG 561
Db 3657 TCTCCAGCTGAAGAGCCACGCCAACCCCAAAATACCTTAAAGAGAGAGGAGGAAG 3716
QY 562 AAAAAATGCTGATCTCT----- 576
Db 3717 AAAAAATGCTGATCTCTCTGTAAGCACTTTTCTTGAATGCTTTCAGACTCGTTTTCAGA 3776
QY 577 -----GAGGAA 582
Db 3777 GGAATTCACAGATCTTCTCTCATGATATATATCTTTTGTGATATGCTTTCACAGAGAA 3836
QY 583 ATAAATCATCAGCTCCCGGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACAGAGT 642
Db 3837 ATAAATCATCAGCTCCCGGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACAGAGT 3896
QY 643 GCTTCTGCGAGAAAGAGAACCGTGTGCTGAGAGAAAGAGAGAGAGCCATCAATGAGCC 702
Db 3897 GCTTCTGCGAGAAAGAGAACCGTGTGCTGAGAGAAAGAGAGAGAGCCATCAATGAGCC 3956
QY 703 AAAAGTTGCAACCAAAACCTTTCTTTCAGAGTGGTTCTCGACCATCTTCTATCTATAC 762
Db 3957 AAAAGTTGCAACCAAAACCTTTCTTTCAGAGTGGTTCTCGACCATCTTCTATCTATAC 4016
```

FEATURES		source	
Location/Qualifiers	1. .79186		
/organism="Arabidopsis thaliana"			
/mol_type="genomic DNA"			
/db_xref="taxon:3702"			
/chromosome="3"			
/clone="K13E13"			
/ecotype="Columbia"			
392. .1558			
/note="unnamed protein product; gb AAC69119.1			
gene id:K13E13.1			
similar to unknown protein			
/codon_start=1			
/evidence=not experimental			
/protein_id="BAB01688.1"			
/db_xref="GI:5280309"			
/translation="MKTIQEQLPNDLVEELICRVFATSLRRLRSTCKAMNRLPKGRI			
LASHKFEKSAQKQFSLRNDYRIFPSFNLHGNPSLELKLSELIDPHSKNSAAPEI			
SRVHCEGLLQSLQDSRVVVMPLNTGTWIRTDGFRQKGRSFDVGYVYQKGRS			
WIKSKLLCYVRGTKEFYDFDSWRILDDIIPRGSISGVSELSKLVGNTWPAK			
WTGEPERTISLLAKFDVTEKSVPLLPYOSRRLFOASSLSVVREDKLSLVLLDDDS			
KTEIWNVIDETTKGAVSWTKVALDLSHLQIQINDGCSFFLGEDKKVVMFCKLIDDS			
NKVDMMVYIGEDNVVTEVGFGVDEMDCRAVLINVVPSLVQIERAGGNRRKH"			
complement (2742. .3977)			
/note="unnamed protein product; gb AAF01602.1			
gene id:K13E13.2			
similar to unknown protein			
/codon_start=1			
/evidence=not experimental			
/protein_id="BAB01689.1"			
/db_xref="GI:5280310"			
/translation="MRFVAPPSPSGDMSPPSSPSSEILSRSSDPPLEFSPPLIA			
MVVLAAFLVITYSRLLISRRFLSPLEPRRWRRCRRRLHLHSSASSASTSSDLRS			
FSPFFDSFHYSSYSPLDSDSVIKTLPLFLYSAACRTGPAVKTSAAACROCAVCL			
LEFEGDYVRLTCLCFHAFLEICDEWLRSHPNCLCTAILGASGVLITPMSFPIFLM			
APRIPLDDENNAIIRGEITPSSNNMTIAADTTNDQEIIRASVEEISSPAISFR			
ELKYSFPECRESSESRVMEPATVSPRYRRSTWNKRSQFPFCNLISKRVSFVRY			
RSTKPPFRSSAGVFPYISERIPATGSSRRRTKSTSPMFRTAPHSSSRLRCGDP			
EALLSPERERDTCRAEM"			
complement (join (5713. .5792, 5883. .6006, 6112. .6349,			
6579. .6697, 6873. .6931, 7161. .7326, 7623. .7682))			
/note="unnamed protein product; gene_id:K13E13.3			
unknown protein"			
/codon_start=1			
/evidence=not experimental			
/protein_id="BAB01690.1"			
/db_xref="GI:5280311"			
/translation="MEFVVEEGKLVNEDCSTLVLPAISIGNVQLAVDLVSSSTGAE			
ENIAFAASSGKKHIIVLSSLDLDFORLNDMSRGPOVYVYLSNAESDGRDDHCRCLFG			
RLHEYDSEGRCKYLVSSVFKEKNSVEELALPSEDELEIDYYPSPILPAALPSAFKARGL			
KMTCLCYCSGEGDNIPAFLLAEASSKLMGLTFDKFHGEEGKQWQIPYISWKSGWIGAP			
DMSMF"			
15532. .16953			
/note="gene_id:K13E13.4"			
/codon_start=1			
/evidence=not experimental			
/product="En/Spm-like transposon protein-like"			
/protein_id="BAB01691.1"			
/db_xref="GI:5280312"			
/translation="MVIATETGSSHHRLTFASFLNSDPAADBEESQSHGVDDIINPS			
SSNNKINQDVHNNVROCSNASTSGESSNOVLSPWNQYTSPPYVTDITPMSDPSPI			
WNQYTSPPYKSPWIYQTRNTDDDTDGLIGTIVROGHVYLSAASGDLITPGSDSKNI			
RWKDLKDHTGKSTSGLVKAIIVTGNRIPTFGQDKIRKVRGSKRRRTGGYSGISGL			
PTLKEFLTGSVNPKNYVEVERKNVLRIKHYDAVCLSNELGLYSGSWDKTLKW			
RLSDSKCLISQIADDAINTVAAGFDLLTGTSGADGTLKWKRELOQKTKTFFLVNL			
MQGNVATLAVNTAAVYVCGSDGTGVNEWQKYLSHGGTTLRGHLRALVLCIAAAGS			
LVLSSGADKNICVWRNGDGSCLSVLMDHVGPVKCLTAVEDDDGEGHREKQDKQIV			
YSGSLDKSVKVRVTSASIVIG"			
19263. .19350			
/product="tRNA-Tyr"			
/note="codon recognized; UAC gene_id:K13E13.5"			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			
CDS			

CDS	/evidence=not_experimental 19930..20860 /note="gb AAD43171.1 gene_id:K13B13.6 similar to unknown protein" /pseudo /codon_start=1 /evidence=not_experimental /note="(21250..21381,21635..21952,22001..22061,22130..22200) /notes="unnamed protein product; gb AAC72857.1 gene_id:K13B13.7 similar to unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB01692.1" /db_xref="GI:9280313" /translation="MYTQNTKARVTSVSHRSQQDPSPVKKFKLVLPSTMKDKRM KIPPRVKLGSKLSEVVTLETPAGFKRSIKLKRIGEEIWFHEGSEFAEASIEGH FLFPEYKENSFRVLIENVSACETKPLDAVHIIDSDDDIIDITKEFDTEHKSRRS RIELDKIILHDLDVMPMHVLEBEEEDKKVFRG" complement (22331..23749) /notes="unnamed protein product; gb AAF15933.1 gene_id:K13B13.8 similar to unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB01693.1" /db_xref="GI:9280314" /translation="NSSVFPGPRFLSLLOONSKTLIQAKIHAQLVINGCHDSLFK LIGHYCSKSTSSSKLAHLVPPRGHDKFLFNTLLKSKPESDIRIPANYASKS LILYNERTFVVGACARSASSALVRIGVHGMVKKGLGYESELIGTLLHIFYAKN GLYLRKFVDEMPERTSVTNAMIGGYCSHKDKGNHNAKAMVLRFRSCGSGVRP TDTTWVCLSAISQTLGSLGIVHGYIEKLGTPEDVDFIGTALVDMYKSCGCLNA FVSFELMKVNTVTSMTATGLANGRNETNLLNMAESGIKPNEITFTSLISAYR HGLVEGIELEKSMKTRGVTPVIEHYCCIVDLGKAGRIQENAVQIFILLAPIKDAL HLSLNCAGSIVGETWGEIEIGKALLEIEREBEKLGSCEDEIVALSINVAHKGWVE VSKLRKEMERIKTRPGYSFV" 25146..26393 /notes="unnamed protein product; gb AAC69119.1 gene_id:K13B13.9 similar to unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB01694.1" /db_xref="GI:9280315" /translation="WTIPICNDLVDLEILCRVPARNLKRSTSKRWNRFLKDDRRFA REHMKAPKEYLPLMTSEVRICPVISINLQGVPSVVLKRELSLDPDYSHQPDGRV FHCDDLVCNVHVKPRYGSKIIVNPLTGQTRWIAZGRWKEIEYRFVIGYCYQDE NDCSCKIKYKILCFYPNGQDTBIYELNYSDRWTRTPDGDLPWTLYISEQTSMNG NLYLFASEKSPHLGVSLLRFPSTKSLCVTLFYQRPRIEILSISAVRGENISLL LQIDFESKTEIIVTWKIDDTTKGAASVTKVLAFDLSPDLQFSEEVNFLDDEKVV AVCCERLPEQSHRYOCRRVKITDKIVILGEDNKVDVSGSEGEATDSLGIQVI LNVAPSLVOIEAGGGKTRGDD" complement (join(27709..27957,28048..28247,28347..28409, 28512..28692)) /notes="unnamed protein product; gb AAD43153.1 gene_id:K13B13.10 similar to unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB01695.1" /db_xref="GI:9280316" /translation="MDSAHNFKRARLFEDLEDAEVIFFSPSVPSLPPESTVPANK GYASAIOTLFTGPVKPEPTPKIKKRGKKKNADPEINSSAPRDDDDPENRSFYE SASARKTVAERERAINAKTPTNPFPRVLRPSVLYRGCMYLPSPGPAEKYLS GISGFIKVLAEKQMPVRLYKAGRAKFSQGWYFTLENNLGEEDVCFVFLRTDFV LKVTAFRVNEY" complement (join(32300..32554,32669..32996,33144..33619)) /codon_start=1 /evidence=not_experimental /product="oxylase-like protein" /protein_id="BAB01696.1"	/db_xref="GI:9280317" /translation="MGELDEAFIQAEHRPNTHLTNSGDFIFSDIEPTIDLSLEDT HDKTAIAKEIAEACRDEWFFQVINGLPSALHRVVKTAAEFNLTTBEKRVKVRDEV NPMGVHDEBHTKNVRDWEIIPDFLQDSTIVPASPEDETELKLNOMPNQSPHRE VCOEYAREVEKLAFLLELVSLISLGLPGRLTGFNEQTSFLRFNHYPCPNDLALG VGRHKDGGATVLAQDSVGLQVSRSDGMIPVKFISDALIINMGCIQVNTDITW SABHRVNVNTSKERFSIPFFPFSPHEANIEPLEEELISEENPPCYKYNWKGKFFVSRNR SDPKKLEVENIQIDHFKA" complement (join(34938..35198,35304..35631,36034..36494))	Query Match 63.6%; Score 653.4; DB 8; Length 79186; Best Local Similarity 77.2%; Pred. No. 2.1e-178; Indels 285; Gaps 3; Matches 969; Conservative 0; Mismatches 1;	QY 58 CTGAGGGTCCACAGATAAGTTTGTGAGTAAATCAAGGATGAGCTTTCGGTTCGTGTGCA 117 DB 28962 CAGAGGGTCCACAGATAAGTTTGTGAGTAAATCAAGGATGAGCTTTCGGTTCGTGTGCA 28903 QY 118 CTCACAGTACCTGATGGTTCATGTTTGGCGGTGATGAGTAAAGGAAAGCTGACAAATAAT 177 DB 28902 CTCACAGTACCTGATGGTTCATGTTTGGCGGTGATGAGTAAAGGAAAGCTGACAAATAAT 28843 QY 178 TGGTTTCAAGATGGTTGGCAAGAGTTTGTGACCGGTACTCCATTGCGATTGCTTATCTT 237 DB 28842 TGGTTTCAAGATGGTTGGCAAGAGTTTGTGACCGGTACTCCATTGCGATTGCTTATCTT 28783 QY 238 TTGATTTTATGATATGAAGAAATCTGCGCTTCAGCGTCTACATTTTCAATTTATCCAC 297 DB 28782 TTGATTTTATGATATGAAGAAATCTGCGCTTCAGCGTCTACATTTTCAATTTATCCAC 28723 QY 298 TCTGAGATCAATTACCATTTCCACCGGTCTCATGGATTCCGCTCACAAACACTTCAAAACGC 357 DB 28722 TCTGAGATCAATTACCATTTCCACCGGTCTCATGGATTCCGCTCACAAACACTTCAAAACGC 28663 QY 358 GCGCGTGTGTTGAAGACCTTGAAGATGAAGTCCGAGGTCTATCTTCTCTTCTCTG 417 DB 28662 GCGCGTGTGTTGAAGACCTTGAAGATGAAGTCCGAGGTCTATCTTCTCTTCTCTG 28603 QY 418 TACCATCACCACTTCTCTGAGTCTACAGTACCAGCAACAAAGGGTATGCTAGTTCAGCC 477 DB 28602 TACCATCACCACTTCTCTGAGTCTACAGTACCAGCAACAAAGGGTATGCTAGTTCAGCC 28543 QY 478 ATCCAAACCTTGTTCATCTGACCAAGTTAA-----506 DB 28542 ATCCAAACCTTGTTCATCTGACCAAGTTAAAGGTATATTTATAACCAACTGATTCCTTT 28483 QY 507-----506 DB 28482 ATCTATCGTGAATTACGGGTCTTATCATTTCTTTGAGTTGATGTTTGATATTTCTCTTA 28423 QY 507-----AGCTGAAGAGCAACCGCAACCCCAAAATACCTTAAAGAGAGAGGAGGAAGAAG 561 DB 28422 TCTCCAGCTGAAGAGCAACCGCAACCCCAAAATACCTTAAAGAGAGAGGAGGAAGAAG 28363 QY 562 AAAAAATGCTGATTCCT-----576 DB 28362 AAAAAATGCTGATTCCTGTTGAAGCACTTTTCTCTTTGAAATGCTTCAGACTCGTTTTT 28303 QY 577-----GAGGAA 582 DB 28302 GGATTCACAGATTTCTTCTCATGATACATATATCTTTTGTATATTTGCTTCTACAGAGAA 28243 QY 583 ATAAACTCATCAGCTCCGCGAGATGATGATCCAGAGAACCGTTTCAAGTTTCTTACGAGAGT 642 DB 28242 ATAAACTCATCAGCTCCGCGAGATGATGATCCAGAGAACCGTTTCAAGTTTCTTACGAGAGT 28183 QY 643 GCTTCTGCGAGAAAGAGAACCGTGTGCTGAGTAAAGAGAAAGAGAGAGAGCCATCAATGAGCC 702 DB 28182 GCTTCTGCGAGAAAGAGAACCGTGTGCTGAGTAAAGAGAAAGAGAGAGAGCCATCAATGAGCC 28123 QY 703 AAAAACTGCAACCAAAACCCCTTTCTTTCAGAGTGGTTCTCGGACCATCTTATCTATAC 762 DB 28122 AAAAACTGCAACCAAAACCCCTTTCTTTCAGAGTGGTTCTCGGACCATCTTATCTATAC 28063
-----	---	---	---	---

QY 763 AGAGTTGCATCAT----- 776  
Db 28062 AGAGTTGCATCATGTAATAAAAAACATCTTAGGAAGACTTAATCTTATCGGTGCTTT 28003  
QY 777 -----GTATCTTCTCTCTGGG 792  
Db 28002 CACTGATCTCTAAAAAAGCCTTCTGTTTCTGTTTCTCTCAACAGTATCTTCTCTGGG 27943  
QY 793 TTTCTGAGAGTACCTTAAGTGGGATCTCCGGGTTCATCAAGTCCAGCTTCCGGAGAAA 852  
Db 27942 TTTCTGAGAGTACCTTAAGTGGGATCTCCGGGTTCATCAAGTCCAGCTTCCGGAGAAA 27883  
QY 853 CAATGGCTGTTCGATGCTCTCAAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGCTAC 912  
Db 27882 CAATGGCTGTTCGATGCTCTCAAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGCTAC 27823  
QY 913 GAATTCATCTAGAGAACATCTTAGGAGAGGAGAGCGTGTGTGTTGAGCTGCTCAGA 972  
Db 27822 GAATTCATCTAGAGAACATCTTAGGAGAGGAGAGCGTGTGTGTTGAGCTGCTCAGA 27763  
QY 973 ACCAGAGATTTCTGTTTGAAGTGCACAGCCCTTTCGAGTCAACGAGTACGCTGAAA 1027  
Db 27762 ACCAGAGATTTCTGTTTGAAGTGCACAGCCCTTTCGAGTCAACGAGTACGCTGAAA 27708

RESULT 10  
AX101027  
LOCUS AX101027 5000 bp DNA linear PAT 10-APR-2001  
DEFINITION Sequence 1 from Patent WO0121822.  
ACCESSION AX101027  
VERSION AX101027.1 GI:13619883  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1  
Dean, C. and Levy, Y. Y.  
METHODS Methods and means for modification of plant flowering  
characteristics  
JOURNAL Patent: WO 0121822-A 1 29-MAR-2001;  
Plant Bioscience Limited (GB)  
FEATURES  
source Location/Qualifiers  
1..5000  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN  
Query Match 63.6%; Score 653; DB 6; Length 5000;  
Best Local Similarity 77.1%; Pred. No. 1.9e-178;  
Matches 968; Conservative 1; Mismatches 1; Indels 285; Gaps 3;  
QY 58 CTGAGGGTCCAGATAGTTGTGATAAATCAAGATGAGCTTTCCGTTCTCTGTGCA 117  
Db 3117 CAGAGGGTCCAGATAGTTGTGATAAATCAAGATGAGCTTTCCGTTCTCTGTGCA 3176  
QY 118 CTCACAGTACCTGATGCTATCTTGGCGTGTAGGACTTAAGAAAGCTGACAAATAAT 177  
Db 3177 CTCACAGTACCTGATGCTATGTTGTCGCTGTAGGACTTAAGAAAGCTGACAAATAAT 3236  
QY 178 TGGTTTCAAGATGGTTGGCAAGATTTGTTGACCGTTACTCCATTCGATTCGTTATCTT 237  
Db 3237 TGGTTTCAAGATGGTTGGCAAGATTTGTTGACCGTTACTCCATTCGATTCGTTATCTT 3296  
QY 238 TTGATTTTATGATATGAAGAAACTCTGCTTTCAGCGTCTACATTTTCAATTTATCCAC 297  
Db 3297 TTGATTTTATGATATGAAGAAACTCTGCTTTCAGCGTCTACATTTTCAATTTATCCAC 3356  
QY 298 TCTGAGATCAATACCATTTCCACCGCTCTCATGGATTCGGTTCACCAACCACTTCAACGC 357

Db 3357 TCTGAGATCAATTACCATTCACCGGTCTCATGGATTCCGCTCAACAACACTTCAACGC 3416  
QY 358 GCCCGTTTGTGTAAGACCTTGAAGATGAAGATGCCGAGGTCACTTTCTTCTTCTGTG 417  
Db 3417 GCCCGTTTGTGTAAGACCTTGAAGATGAAGATGCCGAGGTCACTTTCTTCTTCTGTG 3476  
QY 418 TACCATCAACCACTTCTGAGTCTACAGTACCAAGCAAAAGGATATGCTAGTTCAAGC 477  
Db 3477 TACCATCAACCACTTCTGAGTCTACAGTACCAAGCAAAAGGATATGCTAGTTCAAGC 3536  
QY 478 ATCCAAACCTTCTTCACTGGACCACTAA----- 506  
Db 3537 ATCCAAACCTTCTTCACTGGACCACTTAAAGTGATATTTATAACCAACTGATTCCTTT 3596  
QY 507 ----- 506  
Db 3597 ATCTATCGCTGATTACGGGTCTTATCATTTCTTTGAGGTGATGCTTGATATTTTCTTTA 3656  
QY 507 -----AGCTGAAGAGCCCAACGCCAACCCCAAAATACCTTAAAGAGAGGAGGAGAG 561  
Db 3657 TCTCAGCTGAGAGCCCAACGCCAACCCCAAAATACCTTAAAGAGAGGAGGAGAGAG 3716  
QY 562 AAAAAATGCTGATCTT----- 576  
Db 3717 AAAAAATGCTGATCTTGGTAAGCACTTTTCTCTTTGAAATGCTTCAGACTCGTTTTCAGA 3776  
QY 577 -----GAGAA 582  
Db 3777 GGATTCACAGATTTCTTCTCATGATACATATATCTTTTGATATTTCTTACAGAGAA 3836  
QY 583 ATAAACTCATCAGCTCCGCGAGATGATGATCCAGAGACCGTTTCAAAAGTTCTACGAGAT 642  
Db 3837 ATAAACTCATCAGCTCCGCGAGATGATGATCCAGAGACCGTTTCAAAAGTTCTACGAGAT 3896  
QY 643 GCTTCTCGAGAGAGAGAACCGTCACTGCAAGAGAGAGAGAGAGAGCCATCAATGCAAGC 702  
Db 3897 GCTTCTCGAGAGAGAGAACCGTCACTGCAAGAGAGAGAGAGAGAGCCATCAATGCAAGC 3956  
QY 703 AAAAGTTGCAACCAACCAACCCCTTCTTCAAGATGGTTCTGCGACCATCTCTATCTATAC 762  
Db 3957 AAAAGTTGCAACCAACCAACCCCTTCTTCAAGATGGTTCTGCGACCATCTCTATCTATAC 4016  
QY 763 AGAGTTGCTCATCAT----- 776  
Db 4017 AGAGTTGCTCATCTGTTAATAAAAAACATCTTAGGAAGACTTAATCTTATCGGTGCTT 4076  
QY 777 -----GTATCTTCTTCTGGG 792  
Db 4077 CACTGATCTCTAAAAAAGACCTTCTGTTTCTCTCAACAGTATCTTCTTCTTCTGGG 4136  
QY 793 TTTCTGAGAGTACCTTAAGTGGGATCTCCGGGTTCATCAAGTCCAGCTTGGGAGAAA 852  
Db 4137 TTTCTGAGAGTACCTTAAGTGGGATCTCCGGGTTCATCAAGTCCAGCTTGGGAGAAA 4196  
QY 853 CAATGGCTGTTCGATGCTCTCAAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGCTAC 912  
Db 4197 CAATGGCTGTTCGATGCTCTCAAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGCTAC 4256  
QY 913 GAATTCATCTTAGAGAACAACTTAGGAGAGAGAGAGCGTGTGTGTTTGAAGTCTCTCAGA 972  
Db 4257 GAATTCATCTTAGAGAACAACTTAGGAGAGAGAGAGCGTGTGTGTTTGAAGTCTCTCAGA 4316  
QY 973 ACCAGAGATTTCTGTTTGAAGTGCACAGCCCTTTCGAGTCAACGAGTACGCTGAAA 1027  
Db 4317 ACCAGATTTCTGTTTGAAGTGCACAGCCCTTTCGAGTCAACGAGTACGCTGAAA 4371

RESULT 11  
AY042868  
LOCUS AY042868 1146 bp mRNA linear PLN 04-JUL-2001  
DEFINITION Arabidopsis thaliana Unknown protein (F13P21.8) mRNA, complete cds.  
ACCESSION AY042868  
VERSION AY042868.1 GI:14596160





**KEYWORDS**  
SOURCE  
ORGANISM

**REFERENCE**  
AUTHORS

**TITLE**  
JOURNAL

**COMMENT**

FLI CDNA.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 1146)  
Nguyen, M., Karlin-Neumann, G., Southwick, A., Lam, B., Miranda, M.,  
Palm, C.J., Bowser, L., Jones, T., Ban, J., Carninci, P., Chen, H.,  
Cheuk, R., Chung, M.K., Hayashizaki, Y., Ishida, J., Kamiya, A.,  
Kawai, J., Kim, C., Lin, J., Liu, S.X., Narusaka, M., Pham, P.K.,  
Sakano, H., Sakurai, T., Satou, M., Seki, M., Shim, P., Yamada, K.,  
Shinozaki, K., Ecker, J., Theologis, A. and Davis, R.W.  
Direct Submission  
Submitted (25-JUN-2001) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA  
e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN  
Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PCEC (SSP) Consortium members carried out the  
sequencing and annotation of the RAFL cDNAs: Nguyen, M.,  
Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C.J.,  
Bowser, L., Jones, T., Ban, J., Chen, H., Cheuk, R., Chung, M.K.,  
Kim, C., Lin, J., Liu, S.X., Pham, P.K., Sakano, H., Shim, P.,  
Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.

Nguyen, M., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed  
equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W.  
(SSP/Stanford) contributed equally to this work as PIs.

**FEATURES**  
source  
1..1146  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="1"  
/clone="RAFL04-16-P22"  
/ecotype="Columbia"  
/note="This clone is in pBluescript"  
1..1146  
/gene="F13P21.8"  
177..857  
/gene="F13P21.8"  
/codon\_start=1  
/product="Unknown protein"  
/protein\_id="AAK68808.1"  
/db\_xref="GI:14596161"  
/translation="WQMSRAQNFNRKARLPEDPELKDQKVIYPSIPSTEPVNGY  
GSTDQSPFKESKAEETPKVLKGRKKKNPPEVNSVTPGDDSENRSKEYESASA  
RKTYTAERERAVNAKFTPTNPFYVLRPSLYRGCIMYLPSPGFAEKLISGIS  
FKTLQJGERQWVRCLYKAGRAKFSQGWYFTLENNIGSDVCVFELLRLTRDFVLEIT  
APRVNEYV"

**ORIGIN**  
Query Match 43.1%; Score 442.6; DB 8; Length 1146;  
Best Local Similarity 73.6%; Pred. No. 2.8e-117;  
Matches 647; Conservative 0; Mismatches 194; Indels 38; Gaps 5;  
QY 151 GGACTAAGGAAGCTGACACAAATTTGTTTCAAGATGTTGGCAGAGTTTGTGAC 210  
DB 14 GGACTAAGGAAGCCCAACAAATCTGTTTCAAGACGGTGGCAGAGTTTGTCAAC 73  
QY 211 CGTTACTCCATTCGATTCGTTATCTTTGATTTTATAGATATGAGGAACCTCGCTTC 270  
DB 74 CGTTTCTCATTCGGATTCGT-----TTCAGATACAAGTTACAGTCTACATTTTCAATTT 129  
QY 271 AGCGTCTAATTTTCAATTTATCCCACTCTGAGATCAATTTACCATTTCCACCGGTCTCATG 330

Db 130 ATCTCTCCACACTCTGAGATCAACACC-----ATTCTAGTAGTGAAGCTCTTATGCAATG 185  
QY 331 GATTTCGGCTCACAAACACTT---CAAACGCGCCGTTTGTTCGAAGACCTTGAAGATGAA 387  
Db 186 GATTTCGGCAGAGATCAGTTTCAACAAAGTGTCTGATTTTGAAGATCTTGAATCAAA 245  
QY 388 GATCGCAGGTCACTTTCTTCTTCTGTGTATCCCATCACTTCTCTGAGTCTACAGTA 447  
Db 246 GATGCTAAGGTCAATTATC-----CATCGATCCCTGAATCTACTGAA 287  
QY 448 CCAGCCCAACAAAGGTATCTAGTTTCAAGCATCAACACTTGTCTACTGGACAGTAAA 507  
Db 288 CAGTGAATAAAGTTATGCGGTTTACAGACATCCAAAGCTTTTCAAGAATCTAAA 347  
QY 508 GCTGAAGAGCCAAACGCCAACCCCAAAATACCTTAAAGAGAGAGGAGGAGAGAGAAAAT 567  
Db 348 GCTGAAGA-----AACGCCAAGTACTTTAGAGAGAGAGAGAGAGAGAGAT 398  
QY 568 GCTGATCTTGAGGAAATAAATCATCAGCTCCGCGAGATGATGATCCAGAGAACCGTTCA 627  
Db 399 CCTAATCCCGAGGAGTAACACTTTCACCTCCCGGTGGAGATGACTCAGAGAACCGCTCA 458  
QY 628 AGTTTCTACGAGAGTCTTCTCGGAGAAAGAGAACCGTCTGACAGAGAGAGAGAGAGA 687  
Db 459 AAGTTTCTACGAGAGTCTTCTCTGTAAGAAAGAGAACTGTAACTGCAGAGGAGAGAGA 518  
QY 688 GCATCAATGACGAGCAAAAGCTTTCGAACCAACAAACCCCTTCTTCAGAGTGGTTTCGCGA 747  
Db 519 GCGTCAATGAGCCAAACATTCGAACCAACAAATCTTACTTTAGAGTTGTTTTCGCGA 578  
QY 748 CCATCTATCTATACAGAGTTTGCATCATGTATCTTCTCTGGGTTTCTCGAGAAAGTAC 807  
Db 579 CCATCATATCTATACAGAGTTTGCATCATGTATCTTGGCATCTGGGTTTCTCGAGAAATAC 638  
QY 808 CTAAGTGGATCTCCGGGTTTCATCAAGTCCAGTCTGCGAGAGAAACAATGGCTGTTCGA 867  
Db 639 CTAAGTGGATATCTGGTTTCATCAAGCTCCAGCTCGGTGAGAAAACAATGGCCAGTGAGG 698  
QY 868 TGCTCTTACAAAGCCGGAGAGCCAAATTCAGTCAAGGATGTACGAATTCATCTCTAGAG 927  
Db 699 TGCTCTTACAAAGCAGGAGAGCTAAGTTTACCAAGGATGTATGATGTACACTTCGAG 758  
QY 928 AACAACTTAGGAGAGAGACCTCTGTGTGTTTGAAGTCTCTCAGAACCAAGATTTTCGTT 987  
Db 759 AACAAATATAGCGAGAGAGATCTATGTGTGTTTGAAGTCTCAGAACTCGGATTTTCGTT 818  
QY 988 TTGAAGTACAGACCTTTTCGAGTCAACGAGTACGCTGA 1026  
Db 819 CTCGAAGTCAACGCGCTTTTCGTTCAATGATGTGTGA 857

RESULT 12  
AX505578  
LOCUS  
Sequence 273 from Patent WO0216655.  
DEFINITION  
ACCESSION  
AX505578.1  
VERSION  
AX505578.1  
KEYWORDS  
SOURCE  
ORGANISM  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
HARPER, J.F., KREPS, J., WANG, X. and ZHU, T.  
Stress-regulated genes of plants, transgenic plants containing  
same, and methods of use  
Patent: WO 0216655-A 273 28-FEB-2002;  
The Scripps Research Institute (US); Syngenta Participations AG  
(CH)  
LOCATION/Qualifiers  
1..681  
/organism="Arabidopsis thaliana"



REFERENCE

AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,  
Cook,D., Kim,D., and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (13-AUG-2004) Department of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA

COMMENT

On Aug 3, 2004 this sequence version replaced gi:50540779.  
----- Genome Center  
Center: Department Of Chemistry And Biochemistry  
The University Of Oklahoma  
Center code:UOKNOR

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 17 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 3281: contig of 3281 bp in length  
\* 3282 3381: gap of unknown length  
\* 3282 5850: contig of 2479 bp in length  
\* 5861 5960: gap of unknown length  
\* 5961 9053: contig of 2993 bp in length  
\* 8954 9053: gap of unknown length  
\* 9054 13155: contig of 4102 bp in length  
\* 13156 13255: gap of unknown length  
\* 13256 16225: contig of 2970 bp in length  
\* 16226 16325: gap of unknown length  
\* 16326 21291: contig of 4966 bp in length  
\* 21292 21391: gap of unknown length  
\* 21392 26668: contig of 5277 bp in length  
\* 26669 26768: gap of unknown length  
\* 26769 31867: contig of 5099 bp in length  
\* 31868 31967: gap of unknown length  
\* 31968 38427: contig of 6460 bp in length  
\* 38428 38527: gap of unknown length  
\* 38528 43888: contig of 5361 bp in length  
\* 43889 43988: gap of unknown length  
\* 43989 52130: contig of 8142 bp in length  
\* 52131 52230: gap of unknown length  
\* 52231 59771: contig of 7541 bp in length  
\* 59772 59871: gap of unknown length  
\* 59872 68004: contig of 8133 bp in length  
\* 68005 68104: gap of unknown length  
\* 68105 83623: contig of 15519 bp in length  
\* 83624 83724: gap of unknown length  
\* 83724 98823: contig of 15100 bp in length  
\* 98824 98923: gap of unknown length  
\* 98924 115789: contig of 16866 bp in length  
\* 115790 115889: gap of unknown length  
\* 115890 243777: contig of 127888 bp in length.

FEATURES

source  
1..243777  
Location/Qualifiers  
/organism="Medicago truncatula"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3880"  
/clone="mth2-23j1"  
/clone\_lib="Medicago truncatula BAC library H2"

ORIGIN

Query Match 12.9%; Score 132.6; DB 2; Length 243777;  
Best Local Similarity 60.7%; Pred.No. 9.5e-27;  
Matches 235; Conservative 0; Mismatches 149; Indels 3; Gaps 1;  
QY 30 TTCTCATCATATCCAGAAAAAGCTGTGAGGGTCCAGATAAGTTTGTGAGTAAATT 89  
|||||  
DB 155526 TTGGAACTAATATGATTTGTAATTGCAGAGGCTTCAGATGATTTTATGAGGAATA 155467  
|||||  
QY 90 CAAGGATAGCTTTCCGTTGCTGCTCACTCAAGTACCTGATGGTCATGTTGGCGTGT 149  
|||||

Db 155466 TGGTGGTGATATTCTTCCAACTGTTACTCTCACAGTTCCTGATGGCAGTGTGTGGCGTGT 155407  
QY 150 AGGACTTAAGGAAAGCTGACAAACAAAATTTGGTTTCAAGATGGTTGGCAAGAGTTTCTTGA 209  
Db 155406 AATAATGAAAAGGTAGACAAACAATTTTGGTTCTTGGCGGTGGAATGATTTCTTCA 155347  
QY 210 CCGTTACTCCATTGCGATTGGTTATCTTTTGTGATTTTGTAGATGAAAGGAAACTCTGCTT 269  
Db 155346 AAACCTATTCCATTAGTACTGGATACCTTGTGGTATTCAAAATATGAAGGAAAGTCGCAATTT 155287  
QY 270 CAGCGCTCTACATTTTCAATTTATCCACCTCTGAGATCAATTACCATTCCACCGGTCTCAT 329  
Db 155286 CACTGTTAATAATTTTGTGCTCGCTACTCTGAGATTAATACTATCAATCACCCGCAAAAG 155227  
QY 330 GGATTCGCGTCCAAACCACTT---CAAAGCGCGCGCTTTGTTTGAAGACCTTGAAGATGA 386  
Db 155226 AAGTAATGAAGCGTCTCTCTTTGGAAAGGTCTTACTATTTTCGAAGAATGGAAGATGA 155167  
QY 387 AGATCCGAGGTCATCTTTCTCTTCTTC 413  
Db 155166 AGATTCTGTGAAATCATGGAATCGTC 155140

Search completed: December 30, 2004, 15:53:45  
Job time : 4778 secs

**THIS PAGE BLANK (USPTO)**



CC production of crop plants, where they are able to control the timing of  
CC flowering, the duration of vernalisation required, the optimum  
CC temperature, or even eliminate the need for vernalisation completely. The  
CC present sequence is the VRN1 coding sequence

XX  
SQ  
Sequence 1495 BP: 403 A; 322 C; 312 G; 458 T; 0 U; 0 Other;

Query Match 100.0%; Score 1027; DB 4; Length 1495;

Best Local Similarity 100.0%; Pred. No. 6.1e-312;  
Matches 1027. Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGCCAGCCCTTTCTTCCATAAGTGTGATTTTCTCATCCATATCCCAAGAAAAACGCTCG	60
Db	269	ATGCCAGCCCTTTCTTCCATAAGTGTGATTTTCTCATCCATATCCCAAGAAAAACGCTCG	328
Qy	61	AGGGTCCAGATAAAGTTGTGAGTAAATCAAGAGTAGCTTTCGGTTCGTGTGCATCT	120
Db	329	AGGGTCCAGATAAAGTTGTGAGTAAATCAAGAGTAGCTTTCGGTTCGTGTGCATCT	388
Qy	121	ACAGTACCTGATGTCATGTTTGGCGTGTAGGACTTAAGGAAAGCTGACAAACAAATTTGG	180
Db	389	ACAGTACCTGATGTCATGTTTGGCGTGTAGGACTTAAGGAAAGCTGACAAACAAATTTGG	448
Qy	181	TTTCAAGATGGTTGGCAAGAGTTTCTTGACCGTTACTCCATTTGCAATTTGGTTATCTTTTG	240
Db	449	TTTCAAGATGGTTGGCAAGAGTTTCTTGACCGTTACTCCATTTGCAATTTGGTTATCTTTTG	508
Qy	241	ATTTTTAGATATGAAGGAAACCTCTGCTTCAGCGTCTACATTTTCAATTTATCCACATCT	300
Db	509	ATTTTTAGATATGAAGGAAACCTCTGCTTCAGCGTCTACATTTTCAATTTATCCACATCT	568
Qy	301	GAGATCAATTACCATTTCCACCGGTCTCATGGATTCGGTTCACAAACACATTTCAAAACGCGCC	360
Db	569	GAGATCAATTACCATTTCCACCGGTCTCATGGATTCGGTTCACAAACACATTTCAAAACGCGCC	628
Qy	361	CGTTTGTTTGAAGACCTTCAAGATGAAGATGCGAGGTCATTTTCCCTTCTTCTGTGTAC	420
Db	629	CGTTTGTTTGAAGACCTTCAAGATGAAGATGCGAGGTCATTTTCCCTTCTTCTGTGTAC	688
Qy	421	CCATCACCATCTTCTGAGTCTCAGTACACAGCCAAAGAGGTGTCTAGTTTCAGCCATC	480
Db	689	CCATCACCATCTTCTGAGTCTCAGTACACAGCCAAAGAGGTGTCTAGTTTCAGCCATC	748
Qy	481	CAAACTTGTCTCACTCGACCGAGTTAAAGCTGAAGAGCCAAACGCCAATTAACCTTACCT	540
Db	749	CAAACTTGTCTCACTCGACCGAGTTAAAGCTGAAGAGCCAAACGCCAATTAACCTTACCT	808
Qy	541	AAAAAGAGAGGGAGGAAGAAAAATGCTGATCTTGAGGAAATAAATCATCAGTCTCG	600
Db	809	AAAAAGAGAGGGAGGAAGAAAAATGCTGATCTTGAGGAAATAAATCATCAGTCTCG	868
Qy	601	CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTAGAGAGTGCCTTCGCGAGAAAGAGA	660
Db	869	CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTAGAGAGTGCCTTCGCGAGAAAGAGA	928
Qy	661	ACCGTGACTCGAAGAAAGAGAGAGAGCCATCAATGCAGCCAAAACGTTTCGAACCAACA	720
Db	929	ACCGTGACTCGAAGAAAGAGAGAGAGCCATCAATGCAGCCAAAACGTTTCGAACCAACA	988
Qy	721	AACCTTTTCTTCAGAGTGGTTCTGCGACCATCTCTATCTATACAGAGTTGCATCATGTAT	780
Db	989	AACCTTTTCTTCAGAGTGGTTCTGCGACCATCTCTATCTATACAGAGTTGCATCATGTAT	1040
Qy	781	CTTCCCTTCTGGGTTTCTGAGAGTACTTAAGTGGGATCTCCGGGTTTCATCAAAGTCCAG	840
Db	1049	CTTCCCTTCTGGGTTTCTGAGAGTACTTAAGTGGGATCTCCGGGTTTCATCAAAGTCCAG	1100
Qy	841	CTTCGGAGAAACAAATGGCCCTGTTTCGATGTCTCTCAAAAGCCGGAGAGCCAAATTCAGT	900
Db	1109	CTTCGGAGAAACAAATGGCCCTGTTTCGATGTCTCTCAAAAGCCGGAGAGCCAAATTCAGT	960
Qy	901	CAAGGATGGTACGAATTCATCTTAGAGAAACAACTTAGAGAGAGGAGCGTCTGTGTGTTT	

Db	1169	CRAGGATGGTACGAATTC	CTCTAGAGAAACA	CTTAGGAGAAGGAGACG	TCTGTGTGTTTT	1222
Qy	961	GAGCTGCTCAGACACAGAGATTT	CGTTTTGAAAGTCACAGCC	TTTCGAGTCAACGAGTAC		1020
Db	1229	GAGCTGCTCAGAACCAGAGNTTC	CGTTTTGAAAGTCACAGCC	TTTCGAGTCAACGAGTAC		1288
Qy	1021	GTCTGAA	1027			
Db	1289	GTCTGAA	1295			

## RESULT 2

ADN73922  
ID ADN73922 standard: cDNA; 1026 BP.

XX  
AC ADN73922:

15-JUL-2004 (first entry)

XX ————— dna increased in r2ps/dna expressing plants seqID 1817.

XX plant: transgenic: E2Fa/Dpa transcription factor;

gene; ss; plant; transgenic; bzfr/bfr animal feed product; thale cress; growth regulator

growth regulator; animal feed product; animal cell wall biosynthesis; nitrogen metabolism; carbon metabolism.

XX Arabidongs thaliana.

XX PN WO2004035798-A2.

XX  
29-APR-2004-XX  
PF 20-OCT-2003: 2003WO-EP011658.

18-OCT-2002: 2002EP-00079408.

XX  
PA (CROP-) CROPDESIGN NV.

XX Inze D. De Veylder L, vlieghe K;  
PT

XX  
DR  
WPT: 2004-348466/32.

Altering plant characteristics, useful for producing plants for enzyme or pharmaceutical production comprises modifying in a plant, expression of one or more nucleic acids and/or modifying level or activity of one or more proteins.

XX  
PS  
Claim 1: SEO ID NO 1817; 134pp; English.

This invention relates to a novel method for altering one or more plant characteristics. Specifically, it refers to identifying genes that are up- or down-regulated in transgenic plants overexpressing the heterodimeric E2Fa/DPA transcription factor of Arabidopsis and using these sequences to alter plant characteristics accordingly. The present invention describes generating transgenic plants for the production of growth regulators, enzymes, therapeutics, pharmaceuticals and animal feed products, where the altered plant characteristics are selected from increased yield or biomass, enhanced survival capacity, stress tolerance, plant architecture or physiology, altered endoreplication, biochemistry, signal transduction, storage lipid mobilisation and/or altered photosynthesis, each relative to the corresponding wild type plants. Accordingly, these sequences can also be useful as positive or negative selectable markers during transformation of cells or tissues. The identified genes play a role in a variety of biological processes such as DNA replication, cell wall biosynthesis, nitrogen and/or carbon metabolism or they function as transcription factors. This polynucleotide sequence is thale cress cDNA represented 1.3 fold or more in plants overexpressing the E2Fa/DPA factor given in an exemplification of the invention.

XX  
C- - - - - 1026 ED. 29E A. 235 C. 228 G: 268 T: 0 U: 0 Other;

Query: Match

Query Match 99.5%; Score 100;  
Best Local Similarity 100.0%; Pred. No. 1e-311;

Matches 1026; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGCCAGCGCCCTTCTTCCATAAGTTGATTTTCTCATCTCACTATCCAGAAAAACGCTCG	60
Db	1	ATGCCAGCGCCCTTCTTCCATAAGTTGATTTTCTCATCTCACTATCCAGAAAAACGCTCG	60
QY	61	AGGGTCCAGATAAGTTGTGAGTAAATTCAGAGATGAGCTTTCGGTGTGCTGTGCACTC	120
Db	61	AGGGTCCAGATAAGTTGTGAGTAAATTCAGAGATGAGCTTTCGGTGTGCTGTGCACTC	120
QY	121	ACAGTACCTGATGCTCATGTTTCCGCGTGTAGGACTAAGGAAGCTGACACAAAAATTTGG	180
Db	121	ACAGTACCTGATGCTCATGTTTCCGCGTGTAGGACTAAGGAAGCTGACACAAAAATTTGG	180
QY	181	TTTCAAGATGGTTGGCAAGATTTGTTGACCGTTACTCCATTCGCAATGGTTATCTTTTG	240
Db	181	TTTCAAGATGGTTGGCAAGATTTGTTGACCGTTACTCCATTCGCAATGGTTATCTTTTG	240
QY	241	ATTTTATAGATATGAAGAACTCTGCTTCAGGCTCTACATTTTCAATTTATCCCACTCT	300
Db	241	ATTTTATAGATATGAAGAACTCTGCTTCAGGCTCTACATTTTCAATTTATCCCACTCT	300
QY	301	GAGATCAATTTACCATTTCCACCGTCTCATGATTCGCTCACAAACCACTTCAAAACGGCC	360
Db	301	GAGATCAATTTACCATTTCCACCGTCTCATGATTCGCTCACAAACCACTTCAAAACGGCC	360
QY	361	CGTTGTTTGAAGACCTTGAAGATGAAGATGCGAGGTCTATCTTCTTCTCTGTGTAC	420
Db	361	CGTTGTTTGAAGACCTTGAAGATGAAGATGCGAGGTCTATCTTCTTCTCTGTGTAC	420
QY	421	CCATCACCACTTCTGAGTCTACAGTACCAGCAACAAAGGGTATGCTAGTTCAAGCATC	480
Db	421	CCATCACCACTTCTGAGTCTACAGTACCAGCAACAAAGGGTATGCTAGTTCAAGCATC	480
QY	481	CAAACTTCTTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG	540
Db	481	CAAACTTCTTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG	540
QY	541	AAAAAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	600
Db	541	AAAAAGAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	600
QY	601	CGAGATGATGATCCAGAGAACCTTCAAAGTTCTAAGAGTCTGCTGCGGAGAAAGAGA	660
Db	601	CGAGATGATGATCCAGAGAACCTTCAAAGTTCTAAGAGTCTGCTGCGGAGAAAGAGA	660
QY	661	ACGCTGACTGCAG	720
Db	661	ACGCTGACTGCAG	720
QY	721	AACCTTCTTCTGAGAGTGGTTCTGCGACCACTCTATATACAGAGGTTGCATCATGTAT	780
Db	721	AACCTTCTTCTGAGAGTGGTTCTGCGACCACTCTATATACAGAGGTTGCATCATGTAT	780
QY	781	CTTCTCTTCTGGTTTCTGAGAGTACCTAAGTGGATCTCCGGTTTCATCAAAGTCCAG	840
Db	781	CTTCTCTTCTGGTTTCTGAGAGTACCTAAGTGGATCTCCGGTTTCATCAAAGTCCAG	840
QY	841	CTTGGCGAGAGCAATAGGCTGTTCGATGCTCTACAAAGCCGGAGAGCAAAATTCAGT	900
Db	841	CTTGGCGAGAGCAATAGGCTGTTCGATGCTCTACAAAGCCGGAGAGCAAAATTCAGT	900
QY	901	CAAGGATGGTACGAATTTCACTCTAGAGAACCACTTAGGAGAGAGAGAGAGAGAGAG	960
Db	901	CAAGGATGGTACGAATTTCACTCTAGAGAACCACTTAGGAGAGAGAGAGAGAGAGAG	960
QY	961	GAGCTGCTCAGAACCAAGAGATTTCTGTTTGAAGTGAAGAGAGAGAGAGAGAGAG	1020
Db	961	GAGCTGCTCAGAACCAAGAGATTTCTGTTTGAAGTGAAGAGAGAGAGAGAGAGAG	1020
QY	1021	GTCTGA 1026	
Db	1021	GTCTGA 1026	

RESULT 3	
AAC39629	
ID	AAC39629 standard; DNA; 1510 BP.
XX	
AC	AAC39629;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 25332.
XX	
KW	Hybridisation assay; genetic mapping; gene expression control;
KW	protein identification; signal transduction pathway; metabolic pathway;
KW	promoter; termination sequence; ss.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
XX	25-FEB-2000; 2000EP-00301439.
PF	
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127462P.
PR	06-APR-1999; 99US-0128234P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
PR	23-APR-1999; 99US-0130891P.
PR	28-APR-1999; 99US-0131449P.
PR	30-APR-1999; 99US-0132048P.
PR	30-APR-1999; 99US-0132407P.
PR	04-MAY-1999; 99US-0132484P.
PR	05-MAY-1999; 99US-0132485P.
PR	06-MAY-1999; 99US-0132486P.
PR	06-MAY-1999; 99US-0132487P.
PR	07-MAY-1999; 99US-0132863P.
PR	11-MAY-1999; 99US-0134256P.
PR	14-MAY-1999; 99US-0134218P.
PR	14-MAY-1999; 99US-0134219P.
PR	14-MAY-1999; 99US-0134221P.
PR	14-MAY-1999; 99US-0134370P.
PR	18-MAY-1999; 99US-0134768P.
PR	19-MAY-1999; 99US-0134941P.
PR	20-MAY-1999; 99US-0135124P.
PR	21-MAY-1999; 99US-0135353P.
PR	24-MAY-1999; 99US-0135629P.
PR	25-MAY-1999; 99US-0136021P.
PR	27-MAY-1999; 99US-0136392P.
PR	28-MAY-1999; 99US-0136782P.
PR	01-JUN-1999; 99US-0137222P.
PR	03-JUN-1999; 99US-0137528P.
PR	04-JUN-1999; 99US-0137502P.
PR	07-JUN-1999; 99US-0137724P.
PR	08-JUN-1999; 99US-0138094P.
PR	10-JUN-1999; 99US-0138540P.
PR	10-JUN-1999; 99US-0138847P.
PR	14-JUN-1999; 99US-0139119P.
PR	16-JUN-1999; 99US-0139452P.
PR	16-JUN-1999; 99US-0139453P.
PR	17-JUN-1999; 99US-0139492P.
PR	18-JUN-1999; 99US-0139454P.
PR	18-JUN-1999; 99US-0139455P.
PR	18-JUN-1999; 99US-0139456P.

[illegible]

Db 406 ACAGTACCTGATGGTCATGTTGGCGTGTAGGACTAAGGAAGCTGACAAACAAATTTGG 465  
Qy 181 TTTCAAGATGGTTGGCAAGAGTTTGTGACCGTTACTCCATTGCGATTGGTTATCTTTTG 240  
Db 466 TTTCAAGATGGTTGGCAAGAGTTTGTGACCGTTACTCCATTGCGATTGATTATCTTTTG 525  
Qy 241 ATTTTATAGATGAAGAACTCTGCGCTTCAGGCTCTACATTTTCAATTTATCCCACTCT 300  
Db 526 ATTTTATAGATGAAGAACTCTGCGCTTCAGGCTCTACATTTTCAATTTATCCCACTCT 585  
Qy 301 GAGATCAATTAACATTCACCGTCTCATGGATTCGGCTCACAAACCACTTCAAAACGCGC 360  
Db 586 GAGATCAATTAACATTCACCGTCTCATGGATTCGGCTCACAAACCACTTCAAAACGCGC 645  
Qy 361 CGTTTCTTTGAAGACCTTGAAGATGAAGATGCGGAGTCACTTTCTCTCTCTCTGTGTAC 420  
Db 646 CGTTTCTTTGAAGACCTTGAAGATGAAGATGCGGAGTCACTTTCTCTCTCTGTGTAC 705  
Qy 421 CCATCAACCTCTCTGAGTCTACAGTACAGGCAACAAAGGGTATGCTAGTTCAGCCATC 480  
Db 706 CCATCAACCTCTCTGAGTCTACAGTACAGGCAACAAAGGGTATGCTAGTTCAGCCATC 765  
Qy 481 CAAACCTCTCTGAGTCTACAGTACAGGCAACAAAGGGTATGCTAGTTCAGCCATC 540  
Db 766 CAAACCTCTCTGAGTCTACAGTACAGGCAACAAAGGGTATGCTAGTTCAGCCATC 825  
Qy 541 AAAAAGAGGGAGGAGGAAGAAATGCTGATCTCTGAGGAATAAATCACTCAGCTCCG 600  
Db 826 AAAAAGAGGGAGGAGGAAGAAATGCTGATCTCTGAGGAATAAATCACTCAGCTCCG 885  
Qy 601 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGTGTCTTCGGAAGAAAGAGA 660  
Db 886 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGTGTCTTCGGAAGAAAGAGA 945  
Qy 661 ACCGTGACTGCAGAGAGAGAGAGAGCCATCAATGCGAGCCAAACGTTGCAACCAACA 720  
Db 946 ACCGTGACTGCAGAGAGAGAGAGAGCCATCAATGCGAGCCAAACGTTGCAACCAACA 1005  
Qy 721 AACCTCTCTCAGAGTGGTTCTGCGACCATCTATCTATACAGAGTTGCAATCATGTAT 780  
Db 1006 AACCTCTCTCAGAGTGGTTCTGCGACCATCTATCTATACAGAGTTGCAATCATGTAT 1065  
Qy 781 CTTCTCTCTGGTTCTCAGAGTACCTAAGTGGATCTCCGGGTTCAATCAAGTCCAG 840  
Db 1066 CTTCTCTCTGGTTCTCAGAGTACCTAAGTGGATCTCCGGGTTCAATCAAGTCCAG 1125  
Qy 841 CTTGCGGAGAAACAATGGCTGTTCGATGCTCTACAAAGCGGAGAGCCAAATTCAGT 900  
Db 1126 CTTGCGGAGAAACAATGGCTGTTCGATGCTCTACAAAGCGGAGAGCCAAATTCAGT 1185  
Qy 901 CAAGGATGTTACGAATTCATCTAGAGAACAACTTAGGAGAGGAGACGCTCTGTGTGTT 960  
Db 1186 CAAGGATGTTACGAATTCATCTAGAGAACAACTTAGGAGAGGAGACGCTCTGTGTGTT 1245  
Qy 961 GAGCTGCTCAGAACCAAGATTTCTGTTTGAAGTGCAGAGCTTTCCAGTCAACGAGTAC 1020  
Db 1246 GAGCTGCTCAGAACCAAGATTTCTGTTTGAAGTGCAGAGCTTTCCAGTCAACGAGTAC 1305  
Qy 1021 GTCTGAA 1027  
Db 1306 GTCTGAA 1312

## RESULT 4

AAC51591  
ID AAC51591 standard; DNA; 1509 BP.

XX AC AAC51591;

XX AC AAC51591;

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 69076.

XX

KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
XX promoter; termination sequence; ss.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0128785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 08-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130443P.

XX 23-APR-1999; 99US-0130510P.

XX 28-APR-1999; 99US-0130891P.

XX 30-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 04-MAY-1999; 99US-0132407P.

XX 05-MAY-1999; 99US-0132484P.

XX 06-MAY-1999; 99US-0132485P.

XX 06-MAY-1999; 99US-0132486P.

XX 07-MAY-1999; 99US-0132487P.

XX 11-MAY-1999; 99US-0132863P.

XX 14-MAY-1999; 99US-0134218P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134221P.

XX 18-MAY-1999; 99US-0134370P.

XX 19-MAY-1999; 99US-0134768P.

XX 20-MAY-1999; 99US-0134941P.

XX 21-MAY-1999; 99US-0135124P.

XX 24-MAY-1999; 99US-0135353P.

XX 25-MAY-1999; 99US-0135629P.

XX 27-MAY-1999; 99US-0136021P.

XX 28-MAY-1999; 99US-0136392P.

XX 01-JUN-1999; 99US-0136782P.

XX 03-JUN-1999; 99US-0137222P.

XX 04-JUN-1999; 99US-0137528P.

XX 07-JUN-1999; 99US-0137502P.

XX 08-JUN-1999; 99US-0137724P.

XX 10-JUN-1999; 99US-0138094P.

XX 10-JUN-1999; 99US-0138540P.

XX 14-JUN-1999; 99US-0138847P.

XX 16-JUN-1999; 99US-0139119P.

XX 16-JUN-1999; 99US-0139452P.

XX 17-JUN-1999; 99US-0139453P.

XX 18-JUN-1999; 99US-0139492P.

XX 18-JUN-1999; 99US-0139454P.

XX 18-JUN-1999; 99US-0139455P.

XX 18-JUN-1999; 99US-0139456P.

XX 18-JUN-1999; 99US-0139457P.

XX 18-JUN-1999; 99US-0139458P.

XX 18-JUN-1999; 99US-0139460P.

XX 18-JUN-1999; 99US-0139461P.

XX 18-JUN-1999; 99US-0139462P.

XX 18-JUN-1999; 99US-0139463P.

XX 18-JUN-1999; 99US-0139750P.

XX 21-JUN-1999; 99US-0139763P.

XX 22-JUN-1999; 99US-0139817P.

XX 23-JUN-1999; 99US-0139899P.

XX 23-JUN-1999; 99US-0140353P.

PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0143277P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157533P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159329P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159637P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			

Query Match		98.9%;	Score 1016;	DB 3;	Length 1509;
Best Local Similarity		99.9%;	Pred. No. 1.8e-308;		
Matches 1027;		Conservative	0;	Mismatches	0;
				Indels	1;
				Gaps	1;
Qy	1	ATGCCACGCCCTTTCTTCATAGTTGATTTCTCATCCTATCCACCTATCCAGAAAAAGCTCT-	59		
Db	284	ATGCCACGCCCTTTCTTCATAGTTGATTTCTCATCCTATCCAGAAAAAGCTCTG	343		
Qy	60	GAGGTCCACAGATGTTTGTGAGTAATTCAGATGAGCTTTTCGGTTGCTGTTGCACT	119		
Db	344	GAGGTCCACAGATGTTTGTGAGTAATTCAGATGAGCTTTTCGGTTGCTGTTGCACT	403		
Qy	120	CACAGTACCTGATGCTCATGTTTGGCGTGTAGGACTAAGGAAAGCTGACACAAAATTG	179		
Db	404	CACAGTACCTGATGCTCATGTTTGGCGTGTAGGACTAAGGAAAGCTGACACAAAATTG	463		
Qy	180	GTTCACAGATGTTGGCAAGATTTGTGACCGTTACTCCATTCGATGCTGTTATCTTTT	239		
Db	464	GTTCACAGATGTTGGCAAGATTTGTGACCGTTACTCCATTCGATGCTGTTATCTTTT	523		
Qy	240	GATTTTATGATATGAAGGAAACTCTGCTTCAGGCTCTACATTTTCAATTTATCCACTC	299		
Db	524	GATTTTATGATATGAAGGAAACTCTGCTTCAGGCTCTACATTTTCAATTTATCCACTC	583		
Qy	300	TGAGATCAATTACCATTCACCGGTTTCATCGATTCGCTCACACCACTTCAACCGGC	359		





Db 703 CCTAAAAGAGAGGGAGGAAGAAAATGCTGATCCTCGAGGAAATAAACTCATCAGCT 762  
 Qy 598 CCGCAGATGATGATCCAGAGAACCGTTCAAAGTCTTACGAGAGTCTTCTCGAGAAAG 657  
 Db 763 CCGCAGATGATGATCCAGAGAACCGTTCAAAGTCTTACGAGAGTCTTCTCGAGAAAG 822  
 Qy 658 AGAACGTCGCTCAGAGAAAGAGAGAGAGCCATCAATGCAGCCAAAACGTTTGAACCA 717  
 Db 823 AGAACGTCGCTCAGAGAAAGAGAGAGAGCCATCAATGCAGCCAAAACGTTTGAACCA 882  
 Qy 718 ACAAAACCTTCTTCTCAGAGTGGTCTCGAGACCATCTATCTATACAGAGGTTGCATCATG 777  
 Db 883 ACAAAACCTTCTTCTCAGAGTGGTCTCGAGACCATCTATCTATACAGAGGTTGCATCATG 942  
 Qy 778 TATCTTCTTCTGCGTTCGAGAGTACCTAAGTGGGATCTTCCGCGTTCATCAAGTC 837  
 Db 943 TATCTTCTTCTGCGTTCGAGAGTACCTAAGTGGGATCTTCCGCGTTCATCAAGTC 1002  
 Qy 838 CAGCTTCCGAGAGAAACAAATGGCTGTTCGATGCTCTACAAAGCCGGAGAGCCAAATTC 897  
 Db 1003 CAGCTTCCGAGAGAAACAAATGGCTGTTCGATGCTCTACAAAGCCGGAGAGCCAAATTC 1062  
 Qy 898 AGTCAAGGATGGTACCAATTTCACTCTAGAGAACCACTTAGGAGAGGAGACGCTCTGTG 957  
 Db 1063 AGTCAAGGATGGTACCAATTTCACTCTAGAGAACCACTTAGGAGAGGAGACGCTCTGTG 1122  
 Qy 958 TTTGAGCTCTCAGAACCCAGAGATTTCTTTTGA 993  
 Db 1123 TTTGAGCTCTCAGAACCCAGAGATTTCTTTTGA 1158

RESULT 6

AAF62416  
 ID AAF62416 standard; DNA; 5000 BP.

XX AAF62416;

XX 05-NOV-2001 (first entry)

XX A thaliana VRN1 gene.

XX VRN1; vernalisation; flowering; crop; ds.

XX Arabidopsis thaliana.

XX WO200121822-A1.

XX 29-MAR-2001.

XX 13-SEP-2000; 2000WO-GB003525.

XX 17-SEP-1999; 99GB-00022071.

XX (PLAN-) PLANT BIOSCIENCE LTD.

XX Dean C, Levy YY;

XX WP1; 2001-273467/28.

Novel VRN1 polynucleotide sequence encoding a polypeptide which alters  
 PT vernalization response of plant in which VRN1 nucleic acid is expressed,  
 PT useful for influencing and assessing vernalization phenotype of plants.  
 Claim 4; Page 70-73; 91pp; English.  
 The present invention provides the protein and coding sequences of  
 CC Arabidopsis thaliana VRN1. This protein is capable of altering the  
 CC vernalisation responses of a plant. Also provided are a number of PCR  
 CC primers used to isolate the sequences. The sequences are useful in the  
 CC production of crop plants, where they are able to control the timing of  
 CC flowering, the duration of vernalisation required, the optimum  
 CC temperature, or even eliminate the need for vernalisation completely. The  
 CC present sequence is the VRN1 gene

XX	SQ	Sequence 5000 BP; 1447 A; 871 C; 931 G; 1750 T; 0 U; 1 Other;	
		Query Match 63.6%; Score 653; DB 4; Length 5000;	
		Best Local Similarity 77.1%; Pred. No. 5.7e-194;	
		Matches 968; Conservative 1; Mismatches 1; Indels 285; Gaps 3;	
Qy	58	CTGAGGCTCCAGATAAAGTTTGTGAGTAAATTAAGAGATGAGCTTTCCGGTTGCTGTGCA 117	
Db	3117	CAGAGGCTCCAGATAAAGTTTGTGAGTAAATTAAGAGATGAGCTTTCCGGTTGCTGTGCA 3176	
Qy	118	CTCAGATACCTGATGCTCATGTTGGCGTGTAGACTAAGAAAGCTGACAAATAAT 177	
Db	3177	CTCAGATACCTGATGCTCATGTTGGCGTGTAGACTAAGAAAGCTGACAAATAAT 3236	
Qy	178	TGCTTTCAAGATGGTTGGCAAGATTTGTGACCGTTACTCCATTCGCATTCGGTATCTT 237	
Db	3237	TGCTTTCAAGATGGTTGGCAAGATTTGTGACCGTTACTCCATTCGCATTCGGTATCTT 3296	
Qy	238	TTGATTTTATAGATATGAAGAACTCTGCCTTCAGCGTCTACATTTTCAATTTATCCAC 297	
Db	3297	TTGATTTTATAGATATGAAGAACTCTGCCTTCAGCGTCTACATTTTCAATTTATCCAC 3356	
Qy	298	TCTGAGATCAATTAACATTTCCACCGTCTCATGGATTCGGCTCAACAACCTTCAAAACG 357	
Db	3357	TCTGAGATCAATTAACATTTCCACCGTCTCATGGATTCGGCTCAACAACCTTCAAAACG 3416	
Qy	358	GCCCGTTTGTGTAAGACCTTGAAGATGAAGATGCGAGGTCTATCTTCTTCTTCTGTG 417	
Db	3417	GCCCGTTTGTGTAAGACCTTGAAGATGAAGATGCGAGGTCTATCTTCTTCTTCTGTG 3476	
Qy	418	TACCATCAACACTTCTCTGAGTCTACAGTACCAGCAACAAAGGGTATCTAGTTTCAGCG 477	
Db	3477	TACCATCAACACTTCTCTGAGTCTACAGTACCAGCAACAAAGGGTATCTAGTTTCAGCG 3536	
Qy	478	ATCCAAACCTTGTTCATCTGGACCACTTAA 506	
Db	3537	ATCCAAACCTTGTTCATCTGGACCACTTAAAGGTGATATTTATAACCAACTGATTCCTTT 3596	
Qy	507	----- 506	
Db	3597	ATCTATCTGATTAACGGCTTTATATCTTTTGGAGTTGATGTTGATATTTTCTTTA 3656	
Qy	507	-----AGCTGAAGAGCCAAACGCCAACCCCAAAATACTTAAAGAGAGGAGGAAGAAG 561	
Db	3657	TCTCCAGCTGAAGAGCCAAACGCCAACCCCAAAATACTTAAAGAGAGGAGGAAGAAG 3716	
Qy	562	AAAAATGCTGATCTT----- 576	
Db	3717	AAAAATGCTGATCTTCTGTAAGCACTTTTCTCTTTTGAATGCTTCAGACTCGTTTTCAGA 3776	
Qy	577	-----GAGGA 582	
Db	3777	GGATTCACAGATTTCTCTCATGATACATATATCTTTTGTATATTTGCTTACAGAGGA 3836	
Qy	583	ATAAACTCATCAGCTCCCGGAGATGATGATCCAGAGAACCGTTCAAAGTTCTTACGAGAGT 642	
Db	3837	ATAAACTCATCAGCTCCCGGAGATGATGATCCAGAGAACCGTTCAAAGTTCTTACGAGAGT 3896	
Qy	643	GCTTCTCGAGAGAGAACCGTGTCTCAGAGAAAGAGAGAGAGCCATCAATCGAGCC 702	
Db	3897	GCTTCTCGAGAGAGAACCGTGTCTCAGAGAAAGAGAGAGAGCCATCAATCGAGCC 3956	
Qy	703	AAAAGTTTGAACCAACAAACCTTTCTTTCAGAGTGGTTCTGCGACCATCTTATCTATAC 762	
Db	3957	AAAAGTTTGAACCAACAAACCTTTCTTTCAGAGTGGTTCTGCGACCATCTTATCTATAC 4016	
Qy	763	AGAGTTTGCATCAT----- 776	
Db	4017	AGAGTTTGCATCATGTTAATAAAAAACATCTTAGGAAGACTTAATCTTATCGGTGCTTT 4076	
Qy	777	-----GTATCTTCTTCTTCTGGG 792	

Db 4077 CACTGATCTTAAAGAACGCTTCGTGTTCTCTCTCAACAGTATCTTCTCTGGG 4136  
QY 793 TTTGCTGAGAAGTACTTAAGTGGGATCTCCGGGTTTCATCAAAGTCCAGCTTGGCGAGAA 852  
Db 4137 TTTGCTGAGAAGTACTTAAGTGGGATCTCCGGGTTTCATCAAAGTCCAGCTTGGCGAGAA 4196  
QY 853 CAATGGCTGTGATGTCCTTACAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGGTAC 912  
Db 4197 CAATGGCTGTGATGTCCTTACAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGGTAC 4256  
QY 913 GAAATTCAGTCTAGAGCAACCTTAGGAGAGAGAGAGCTCTGTGTTGTAGCTCTCTCAGA 972  
Db 4257 GAAATTCAGTCTAGAGCAACCTTAGGAGAGAGAGAGCTCTGTGTTGTAGCTCTCTCAGA 4316  
QY 973 ACCAGAGATTTGTTTTGAAAGTGAACGCTTTTCGAGTCAACGAGTACGTCGAA 1027  
Db 4317 ACCAGAGATTTGTTTTGAAAGTGAACGCTTTTCGAGTCAACGAGTACGTCGAA 4371  
  
RESULT 7  
AAC37903  
ID AAC37903 standard; DNA; 753 BP.  
XX AC AAC37903;  
XX  
XX 17-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 19075.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX  
XX Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
PD  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000BP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.



PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 24-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142377P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144006P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 22-JUL-1999; 99US-0145193P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149436P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.

PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.  
Query Match 42.9%; Score 441; DB 3; Length 1394;  
Best Local Similarity 73.5%; Pred. No. 1.5e-127;  
Matches 646; Conservative 0; Mismatches 195; Indels 38; Gaps 5;  
QY 151 GCACTAAGAAAGCTGACAAACAAATTTGGTTTCAAGATGGTTGGCAAGAGTTTGTGAC 210  
DB 278 GGACTAAGAAAGCCAAACAAATCTGTTTCAAGACGGTTGGCAGGAGTTGTCAAC 337  
QY 211 CGTTACTCCATTCGGATTCGTTTATCTTTTGGTTTATAGATATGAAGAACTCTGCCCTC 270  
DB 338 CGTTTCTCCATTCGGATTCGTTTATCTTTTGGTTTATAGATATGAAGAACTCTGCCCTC 393  
QY 271 AGCGTCTACATTTCAATTTATCCACTCTGAGATCAATTTACATTCACCGGTCTCATG 330  
DB 394 ATTCCTCCACTCTGAGATCAACACC-----ATTCTAGTAGTGAAGCTTTATGCAATG 449  
QY 331 GATTCGGCTCAACACCACTTT---CAAACGCGCCGTTTGGTTTGAAGACCTTGAAGATGAA 387  
DB 450 GATTCGGCACAGATCAGTTTCAACAAACGCTGCTCGATTTGTTGAAGATCCTGAACTCAA 509  
QY 388 GATCGGAGGTATCTTTCTTCTCTGTGTACCCATCACCACCTTCTGTGAGTCTACAGTA 447  
DB 510 GATGTAAGGTCAATTTATC-----CATCGAACCTTGAATCTACTGAA 551  
QY 448 CCAGCCCAACAAAGGGTATCTAGTTTCAGCCATCCAAACCTTGTCTACTGGACCAGTTAA 507  
DB 552 CCAATGTAATGAAGTTATGCGCGTTCTACAGCATCCAAAGCTTTTCAAAGATCTAAA 611  
QY 508 GCTGAAGAGCCCAACGCCAAACCCCAAAATATCTTAAAGAGAGGGAGGAAGAAAGAAAT 567  
DB 612 GCTGAAGA-----AACGCCCAAGGTACTTTAAGAGAGAGGAAGGAAGAAAGAAAT 662  
QY 568 GCTGATCTGAGGAATAAATCACTCAGCTCCGCGAGATGATGATCCAGAGAACCGTTCA 627  
DB 663 CCTAATCCCGAGGAAGTAAATCTTTCACTCCCGGTGGAGATGACTCAGAGAACCGCTCA 722  
QY 628 AAGTTCTACGAGAGTCTTTCTGCGAGAAAGAGAACCGTGCACCTGCAGAGAAAGAGAGAGA 687  
DB 723 AAGTTCTACGAGAGTCTTTCTGCTAGAAAGAGAACTGTAACTGCAGAGAAAGAGAGAGA 782  
QY 688 GCCATCAATGACGCCAAACGTTTCGAACCAACAAACCCCTTTTCTTCAGAGTGTGTTCTGCGA 747

DB 783 GCCGTCAATGACGCCAAACAAATTCGAACCAACAAATCCTTACTTTTAGAGTTGTTCTGCGA 842  
QY 748 CCATCCTATCTATACAGAGGTTGCAATCATGTATCTTCTCTGGGTTTCTCAGAAATAC 807  
DB 843 CCATCATATCTATACAGAGGTTGCAATCATGTATCTTCTCAGAAATAC 902  
QY 808 CTAAGTGGGATCTCCGGTTTCATCAAAGTCCAGCTTGGCGGAGAAACAAATGGCCTGTTCGA 867  
DB 903 CTAAGTGGGATATCTGGTTTCATCAAGCTCCAGCTCGGTGAGAAACAAATGGCCAGTGAGG 962  
QY 868 TGTCTCTACAAGCCGGGAGAGCCAAATTCAGTCAAGGATGCTACGAATTCATCTAGAG 927  
DB 963 TGCCTCTACAAGACGGGAGAGCTAAGTTTAGCCAAGGATGTTATGATGAGTTCACACTCGAG 1022  
QY 928 AACCAACTTAGGAGAGGAGACGTCTGTGTGTTTGTAGCTGCTCAGAACCAAGATTTTCGTT 987  
DB 1023 AACATATAGGCAAGGAGATGATGTGTTTGTAGCTACTCAGAACTCGGATTTTCGTT 1082  
QY 988 TTGAAGTGCACACCGCTTTTCGAGTCAACGAGTACGTCGA 1026  
DB 1083 CTCGAAGTCACCGCTTTTCGTCATGATGATGTCGA 1121  
RESULT 9  
AAC53657  
ID AAC53657 standard; DNA; 1396 BP.  
AC AAC53657;  
XX  
XX 18-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 75262.  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.

PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135622P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138542P.
PR	14-JUN-1999;	99US-0138847P.
PR	16-JUN-1999;	99US-0139119P.
PR	17-JUN-1999;	99US-0139452P.
PR	18-JUN-1999;	99US-0139453P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142377P.
PR	13-JUL-1999;	99US-0143342P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	30-AUG-1999;	99US-0151080P.
PR	31-AUG-1999;	99US-0151303P.
PR	01-SEP-1999;	99US-0151304P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155119P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-015753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.

PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161922P.	
PR	28-OCT-1999;	99US-0161933P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match			42.9%; Score 441; DB 3; Length 1396;
Best Local Similarity			73.5%; Pred. No. 1.5e-127;
Matches 646; Conservative 0; Mismatches 195; Indels 38; Gaps 5;			
QY	151	GGACTAAGGAAGCTGACAAACAATAATTTGGTTTCAAGATGGTTGGCAAGAGTTTGTGAC	210
DB	279	GGACTAAGGAAGCCAAACAATAATCTGGTTTCAAGACGGTTGGCAGGAGTTTGTCAAC	338
QY	211	CGTTACTCCATTCGCATTCGTTATCTTTGATTTTAGATATGAAGGAACCTCGCTTC	270
DB	339	CGTTCTCCATTCGGATTTGGT-----TTGATATCAAAAGTTACGTCTACATTTTCAATTT	394
QY	271	AGCGTCTACATTTTCAATTTATCCCACTCTGAGATCAATTTACCATTCCACCGGTCTCATG	330
DB	395	ATCCTCCACACTCTGAGATCAACACC-----ATTCTAGTAGTGAAGCTCTTATGCAATG	450
QY	331	GATTCGCTCACAAACACTT-----CAAAAGCGCCGTTTGTGTTGAAGACCTTGAAGATGAA	387
DB	451	GATTCGCGCAGAAATCAGTTTCAACAAACGTCGATTTGTTGAAGATCCTGAACCTCAA	510
QY	388	GATCGCGAGGTCTATCTTTCTTCTCTGTGTACCCATCACCATTCTCTGAGTCTACAGTA	447
DB	511	GATGCTAAGGTCAATTTATC-----CATCGAACCCCTGAATCTACTGAA	552
QY	448	CCAGCCAAACAAGGGTATGCTAGTTCAGGCATCCAAACCTTGTCTACTGGACCAGTTAA	507
DB	553	CCAGTGAATTAAGTTTATGGCGTTCTACAGCATCCAAAGCTTTTCAAGATCTAAA	612
QY	508	GCTGAAGCCACACGCCAACCCCAAAATACCTTAAAGAGAGGGAGGAAGAAATAAT	567
DB	613	GCTGAAGA-----AAGCCCAAGGTACTTAAGAGAGAGGAAGGAAGAAGAAAT	663
QY	568	GCTGATCTTGAGGAATTAATCATCAGTCCGCGAGATGATGATCCAGAGAACCGTTCA	627
DB	664	CTTAATCCCGAGGAAGTAATCTTCAATCCCGGTGGAGATGACTCAGAGAACCGCTCA	723
QY	628	AAGTTCTACGAGAGTCTTCTGCGAAGAGAGAACCGTCTGACCTGCAAGAAAGAGAGAGA	687
DB	724	AGTTTCTACGAGAGTCTTCTGCTAGAAAGAGAACTGTAACTGCAGAGAAAGAGAGAGA	783
QY	688	GCCATCAATGACGCCAAAACGTTTCGAACCAACAAACCCCTTTCTTCAGAGTGGTTCTGCGA	747
DB	784	GCCGTCAATGCGAGCCAAAACATTCGAACCAACAAATCCTTACTTTAGAGTTGTTCTGCGA	843
QY	748	CCATCCTATCTATACAGAGTTGTCATCATGTATCTTCTCTCTGGGTTTCTGAGAGATAC	807
DB	844	CCATCATATCTATACAGAGTTGTCATCATGTATCTTCTCTCTGGGTTTCTGAGAGATAC	903
QY	808	CTAAGTGGATCTCCGGGTTTCATCAAGTCCAGCTTTCGCGAGAAACAATGGCTGTTCGA	867
DB	904	CTAAGTGGATATCTGGTTTTCATCAAGTCCAGCTCGGTGAGAAACATGGCCAGTGAGG	963
QY	868	TGTCTCTACAAAGCCGGAGAGCCAAATTCAGTCAAGGATGGTACGAATTTCACTCTAGAG	927
DB	964	TGCCTCTACAAAGCAGGAGAGCTAAGTTTAGCCAGGATGGTATGAGTTTCACTCTGAG	1023
QY	928	AACAACCTTAGGAGAGGAGAGCTCTGTGTGTTTGTAGCTCTCAGAACCCAGATTTCTGTT	987
DB	1024	AACAATATAGGCGAGGAGAGATGTGTGTGTTTGTAGCTACTCAGAACTCGGATTTCTGTT	1083
QY	988	TTGAAAGTACAGCCCTTTTCGAGTCAACGAGTAGTCTGA	1026
DB	1084	CTCGAGTACCCGCCCTTTTCGTGTCAATGATGATGTGTA	1122

RESULT 10	
AAC37617	
ID	AAC37617 standard; DNA; 1344 BP.
XX	
AC	AAC37617;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 18041.
XX	
KW	Hybridisation assay; genetic mapping; gene expression control;
KW	protein identification; signal transduction pathway; metabolic pathway;
KW	promoter; termination sequence; ss.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.
PR	01-APR-1999; 99US-0127462P.
PR	06-APR-1999; 99US-0128234P.
PR	08-APR-1999; 99US-0128714P.
PR	16-APR-1999; 99US-0129845P.
PR	19-APR-1999; 99US-0130077P.
PR	21-APR-1999; 99US-0130449P.
PR	23-APR-1999; 99US-0130510P.
PR	23-APR-1999; 99US-0130891P.
PR	28-APR-1999; 99US-0131449P.
PR	30-APR-1999; 99US-0132048P.
PR	30-APR-1999; 99US-0132407P.
PR	04-MAY-1999; 99US-0132484P.
PR	05-MAY-1999; 99US-0132485P.
PR	06-MAY-1999; 99US-0132486P.
PR	06-MAY-1999; 99US-0132487P.
PR	07-MAY-1999; 99US-0132863P.
PR	11-MAY-1999; 99US-0134256P.
PR	14-MAY-1999; 99US-0134218P.
PR	14-MAY-1999; 99US-0134219P.
PR	14-MAY-1999; 99US-0134221P.
PR	18-MAY-1999; 99US-0134370P.
PR	19-MAY-1999; 99US-0134768P.
PR	19-MAY-1999; 99US-0134941P.
PR	20-MAY-1999; 99US-0135124P.
PR	21-MAY-1999; 99US-0135353P.
PR	24-MAY-1999; 99US-0135629P.
PR	25-MAY-1999; 99US-0136021P.
PR	27-MAY-1999; 99US-0136392P.
PR	28-MAY-1999; 99US-0136782P.
PR	01-JUN-1999; 99US-0137222P.
PR	03-JUN-1999; 99US-0137528P.
PR	04-JUN-1999; 99US-0137502P.
PR	07-JUN-1999; 99US-0137724P.
PR	08-JUN-1999; 99US-0138094P.
PR	10-JUN-1999; 99US-0138540P.
PR	10-JUN-1999; 99US-0138847P.
PR	14-JUN-1999; 99US-0139119P.
PR	16-JUN-1999; 99US-0139452P.
PR	16-JUN-1999; 99US-0139453P.
PR	17-JUN-1999; 99US-0139492P.
PR	18-JUN-1999; 99US-0139454P.
PR	18-JUN-1999; 99US-0139455P.
PR	18-JUN-1999; 99US-0139456P.
PR	18-JUN-1999; 99US-0139457P.
PR	18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 21-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147202P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.

PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150586P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 42.8%; Score 439.4; DB 3; Length 1344;

Best Local Similarity 73.4%; Pred. No. 4.7e-127;

Matches 645; Conservative 0; Mismatches 196; Indels 38; Gaps 5;

QY 151 GGACTAAGGAAAGCTGACAAACAAAATTTGGTTTCAAGATGGTTGGCAAGAGTTTGTTCAC 210

Db 226 GGACTAAGGAAAGCCCAACAAAATCTGGTTTCAAGACGGTTGGCAGGAGTTTGTCAAC 285

QY 211 CGTACTCCATTCGCAATGGTTATCTTTTGATTTTATAGATGAAGAACTCGCTTC 270

Db 286 CGTTTCTCCATTCGGAATGGT----TTCAGATACAAAGTTACAGTCTACATTTTCAATTT 341

QY 271 AGCGTCTACATTTTCAATTTATCCCACTCTGAGATCAATTACCACTTCCACCGGTCTCATG 330

Db 342 ATCTCCCACTCTGAGATCAACACC-----ATTCTAGTAGAGTCTTATGCAATG 397

331 GATTTCGCTCACAACCACTT---CAAAACGCGCCGCTTGTGTTGAAGACCTTGAAGATGAA 387  
398 GATTTCGCGCAGAAATCAGTTTCAACAAACGTCGCTGATTTGTTGAAGATCCTGAACCTCAAA 457  
388 GATGCGGAGGTCTATCTTCTCTTCTCTGTGTATCCCATCACCCTTCTCTGAGTCTACAGTA 447  
458 GATGCTAAGGTCTATTATC-----CATCGAACCTGATTTTACTGAA 499  
448 CAGCCCAACAAAGGGTATGCTAGTTTCAGCCATCAAAACCTTGTTCATCGACAGTTAAA 507  
500 CCACTGAATAAAGGTTATGCGGTTCTACAGCCATCAAAAGCTTTTCAAGAAATCTAAA 559  
508 GCTGAGAGCCACGCAACCCCAAAATACCTTAAAGAGAGGAGGAGGAAGAAAT 567  
560 GCTGAAGA-----AAGCGCCAAGGTACTTAAAGAGAGGAGGAAGAAAGAAAT 610  
568 GCTGATCTGAGGAATAAATCATCAGCTCCGAGATGATGATCATCAGAGAACGTTCA 627  
611 CTTAATCCCGAGGAGTAATCTTCACTCCGCTGGAGATGACTCAGAGAACCGCTCA 670  
628 AAGTTCTACGAGAGTCTTCTGCGAGAAAGAGAACCGTCACTGCAAGAAAGAGAGAGA 687  
671 AAGTTCTACGAGAGTCTTCTGCTAGAAAGAGAACTGTAATGCAAGAGAAAGAGAGAGA 730  
688 GCCATCAATGCGCAAAAGCTTCGAACCAACCAACCCCTTCTTCAAGTGGTTCTGCGA 747  
731 GCGGTCAATGCGAGCCAAAACATTCGAACCAACCAATCTTACTTTAGAGTTGTTCTGCGA 790  
748 CCATCTCTATCTATACAGAGGTGTCATCATGTATCTTCTCTGCGTTCTGAGAGTAC 807  
791 CCATCATATCTATACAGAGTTGTCATCATGTTCTGCCATCTGGGTTCTGAGAAATAC 850  
808 CTAAGTGGATCTCCGGGTTCTCAAAAGTCCAGTTCGCGAGAAACAAATGGCTGTTCGA 867  
851 CTAAGTGGATATCTGTTTCTATCAAGCTCCAGCTCGGTGAGAAACAAATGGCCAGTGAGG 910  
868 TGCTCTACAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGGTACGAATTCATCTAGAG 927  
911 TGCTCTACAAAGCAGGAGAGCTAAGTTTAGCAAGGATGGTATGAGTTCACTCGAG 970  
928 AACAATAGGAGAGAGAGAGCTGTGTGTTGAGCTGCTCAGAACAGAGATTTCTGTT 987  
971 AACAATAGGAGAGAGATGATGTGTGTGTGAGCTACTCAGAACTCGGATTTCTGTT 1030  
988 TTGAAGTGCACGCTTTTCGAGTCAACGAGTACGTTCTGA 1026  
1031 CTCGAGTCAACGCTTTTCGTTGTCATGATGTGTGA 1069

## RESULT 11

ABZ12468

ID ABZ12468 standard; DNA; 681 BP.

XX AC ABZ12468;

XX DT 21-JAN-2003 (first entry)

XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 273.

XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX OS Arabidopsis thaliana.

XX PN WO200216655-A2.

XX PD 28-FEB-2002.

XX PF 24-AUG-2001; 2001WO-US026685.

XX PR 24-AUG-2000; 2000US-0227866P.

XX PR 26-JAN-2001; 2001US-0264647P.

XX PR 22-JUN-2001; 2001US-0300111P.

XX XX

PA (SCRI) SCRIPPS RES INST.  
PA (SYGN) SYNGENTA PARTICIPATIONS AG.  
XX Harper JP, Krops J, Wang X, Zhu T;  
XX WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and  
PT producing plants with increased tolerance to these abiotic stresses.

PS Claim 144; SEQ ID NO 273; 577pp + Sequence Listing; English.

CC The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid  
CC representative of expressed polynucleotides in the plant cell with an  
CC array or probes representative of the plant cell genome; and (b)  
CC detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
CC in methods of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office

SQ Sequence 681 BP; 217 A; 139 C; 162 G; 163 T; 0 U; 0 Other;

Query Match 38.7%; Score 397.6; DB 6; Length 681;

Best Local Similarity 76.6%; Pred. No. 4.6e-114; Mismatches 134; Gaps 3;

Matches 538; Conservative 0; Indels 30;

QY 328 ATGATTCGCTCACAACCACTT---CAAAACGCGCCGCTTGTGTTGAAGACCTTGAAGAT 384  
Db 7 ATGATTCGCGACAGATCAGTTCAACAAACGTCGATTTGTTGAAGATCCTGAATC 66  
QY 385 GAAGATGCGAGGTCTTCTTCTTCTGTGTACCCATCACCCTTCTCTGAGTCTACA 444  
Db 67 AAAGATGTAGGTCTATTATCCATCGAAC-----CCTGAATCTACT 108  
QY 445 GTACAGCCCAACAAAGGTATGTTTACAGCCATCCAAACCTTCTCTACTGACAGTT 504  
Db 109 GAACAGTGAATAAAGGTATGCGCGTTCTACAGCCATCCAAAGCTTTTCAAAGATCT 168  
QY 505 AAAGCTGAAGAGCCAAACCCCAACCAAAATAATACCTAAAGAGAGGAGGAGAGAGAA 564  
Db 169 AAAGCTGAAGA-----AACGCGCCAGGTACTTTAAGAGAGAGGAGAGAGAGAG 219  
QY 565 AATGCTGATCTGAGGAATAAATCACTCAGTCTCGCGAGATGATGATCCAGAGAACCTG 624  
Db 220 AATCCTAATCCGAGGAAGTAACTCTTCAACTCCCGGTGGAGATGACTCAGAGAACCCG 279  
QY 625 TCAAGTTCTACGAGAGTCTTCTGCGAGAGAGAACCGTACCTGCGAGAGAGAGAGAG 684  
Db 280 TCAAGTTCTACGAGAGTCTTCTGCTAGAGAGAGAACTGTAACTGCGAGAGAGAGAG 339  
QY 685 AGAGCCATCAATGAGCCAAACACGTTCCGAAACCAACAAACCTTTCTTCTGAGTGTCTG 744  
Db 340 AGAGCCGTCATGAGCCCAAAACATTCGAACCAACAAATCTTACTTTAGAGTTGTTCTG 399  
QY 745 CGACCATCTATCTATACAGAGTTGCAATCATGATCTTCTTCTTCTGCGGTCTGAGAG 804  
Db 400 CGACCATCTATCTATACAGAGTTGCAATCATGATCTTCTTCTGCGGTCTGAGAGAA 459  
QY 805 TACCTAAGTGGGATCTCCGGTTCTATCAAGTCCAGCTTGGCGAGAGAAACAAATGGCTGTT 864  
Db 460 TACCTAAGTGGGATCTGCGTTCTATCAAGCTCCAGCTCGGTGAGAGAAACAAATGGC 519  
QY 865 CGATGCTCTACAAAGCCGGAGAGCCAAATTTCACTCAAGGATGATGATGATTTCACTCTA 924  
Db 520 AGGTGCTCTACAAAGCAGGAGAGCTAAGTTTAGCCAAAGGATGATGATGATTTCACTC 579  
QY 925 GAGAACAACTTAGGAGAGAGAGAGCTGTGTGTGTTTGTAGTCTCTCAGAACCCAGATTTC 984

Db 580 GAGAACATATAGCGAAGGAGATGATGTGTTGGAGTACTCAGAACTCGGGATTTC 639  
QY 985 GTTTCGAAGTGAAGCCCTTCGAGTCAACGAGTACGCTGA 1026  
Db 640 GTTCTCGAAGTCACGCGCTTCGTCGTCATGAGTATGTGA 681

RESULT 12  
ABX62122/c  
ID ABX62122 standard; DNA; 426 BP.  
XX  
AC ABX62122;  
XX  
DT 25-FEB-2003 (first entry)  
XX  
DE Arabidopsis thaliana expressed sequence related polynucleotide #237.  
XX  
KW Transgenic plant; plant; genetically modified cell; environmental stress;  
KW ribozyme creation; disease resistance; stress tolerance;  
KW fungicide screening; insecticide screening; gene; ds.  
XX  
OS Unidentified.  
XX  
PN US2002040490-A1.  
XX  
PD 04-APR-2002.  
XX  
PF 26-JAN-2001; 2001US-00770423.  
XX  
PR 27-JAN-2000; 2000US-0178512P.  
XX  
PA (GORL/) GORLACH J.  
PA (ANY/) AN Y.  
PA (HAMI/) HAMILTON C M.  
PA (PRIC/) PRICE J L.  
PA (RAIN/) RAINES T M.  
PA (YUY/) YU Y.  
PA (RAME/) RAMEKA J G.  
PA (PAGE/) PAGE A.  
PA (MATH/) MATHAW A V.  
PA (LEDF/) LEDFORD B L.  
PA (WOES/) WOESSNER J P.  
PA (HAAS/) HAAS W D.  
PA (GARC/) GARCIA C A.  
PA (KRIC/) KRICKER M.  
PA (SLAT/) SLATER T.  
PA (DAVI/) DAVIS K R.  
PA (ALLE/) ALLEN K.  
PA (HOFF/) HOFFMAN N.  
PA (HURB/) HURBAN P.  
XX  
PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;  
PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
PI Hurban P;  
XX  
WPI; 2003-110411/10.  
XX  
XX Novel Arabidopsis thaliana nucleic acid useful for identifying homologous  
PT or related genes, and to create genetically modified and transgenic  
PT organisms, such as plant cells and plants.  
XX  
PS Claim 1; SEQ ID NO 237; 43pp; English.  
XX  
CC The invention describes an Arabidopsis thaliana nucleic acid (I). The  
CC polypeptide (II) encoded by (I), transgenic plant (III) or genetically  
CC modified cell (IV) are useful for screening a candidate agent for its  
CC biological effect, by combining the candidate agent with (II), (III) or  
CC (IV), and determining the effect of the candidate agent on (II), (III) or  
CC (IV). (I) is useful for identifying homologous or related genes, for  
CC producing compositions that modulate the expression or function of its  
CC encoded protein, for mapping functional regions of the protein, in  
CC diagnosis, for studying associated physiological pathways, for genetic

CC manipulation of cells, preferably plant cells, in screening assays of  
CC various plant strains to determine the strains that are capable of  
CC withstanding a particular disease or environmental stress, for enhancing  
CC or inhibiting production of biosynthetic product in a plant, for  
CC producing polypeptides, as probes for the detection of mRNA in biological  
CC samples, to generate additional copies of (I), to generate ribozymes or  
CC oligonucleotides, as single stranded DNA probes or as triple-strand  
CC forming oligonucleotides, and to create genetically modified and  
CC transgenic organisms, such as plant cells and plants. (II) or (III) is  
CC useful for introducing or improving disease resistance and stress  
CC tolerance in plants, screening biological active agents, e.g.,  
CC fungicides, insecticides, etc., and for elucidating biochemical pathways.  
CC (III) is useful as crops for their enhanced disease resistance, enhanced  
CC traits of interest, for screening programs, as crops which exhibit  
CC enhanced tolerance to environmental stress, or to produce a factor. This  
CC sequence represents a nucleic acid that may correspond to naturally  
CC occurring Arabidopsis thaliana expressed sequences. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from the US patent office at  
CC seqdata.uspto.gov/sequence.html?docID=99990770423  
XX  
SQ Sequence 426 BP; 127 A; 101 C; 82 G; 114 T; 0 U; 2 Other;

Query Match 28.8%; Score 295.4; DB 8; Length 426;  
Best Local Similarity 99.0%; Pred. No. 5.2e-82;  
Matches 296; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 729 CTTGAGAGTGGTCTCGGACCATCTATCTATACAGAGGTTGCATCATGTATCTTCCTTC 788  
Db 426 CTTGAGAGTGGTCTCGGACCATCTATCTATACAGAGGTTGCATCATGTATCTTCCTTC 367  
QY 789 TGGGTTTGTCTGAGAGTACCTTAAGTGGGATCTCCGGTTTCATCAAGTCCAGTTCGGA 848  
Db 366 TGGGTTTGTCTGAGAGTACCTTAAGTGGGATCTCCGGTTTCATCAAGTCCAGTTCGGA 307  
QY 849 GAAACATATGCGCTGTTCGATGTCTCTACAAAGCCGGGAGAGCCAAATTCAGTCAAGGATG 908  
Db 306 GAAACATATGCGCTGTTCGATGTCTCTACAAAGCCGGGAGAGCCAAATTCAGTCAAGGATG 247  
QY 909 GTACGAATTCACCTCTAGAGAACAACTTAGGAGAGGAGACGCTGTGTGTTTGTAGCTGCT 968  
Db 246 GTACGAATTCACCTCTAGAGAACAACTTAGGAGAGGAGACGCTGTGTGTTTGTAGCTGCT 187  
QY 969 CAGAACCCAGAGATTTGTTTGAAGTGACAGCCCTTTGAGTCAACGAGTACGCTGAA 1027  
Db 186 CAGAACCCAGAGATTTGTTTGAAGTGACAGCCCTTTGAGTCAACGAGTACGCTGAA 128

RESULT 13  
ABZ15168  
ID ABZ15168 standard; DNA; 2000 BP.  
XX  
AC ABZ15168;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2973.  
XX  
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200216655-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 24-AUG-2001; 2001WO-US026685.  
XX  
PR 24-AUG-2000; 2000US-0227866P.  
XX  
PR 26-JAN-2001; 2001US-0264647P.  
XX  
PR 22-JUN-2001; 2001US-0300111P.  
XX  
PA (SCRI ) SCRIPPS RES INST.



PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 18-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 21-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160815P.
PR 22-JUL-1999;	99US-0145192P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160989P.
PR 23-JUL-1999;	99US-0145224P.	PR 25-OCT-1999;	99US-0161404P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-014918P.	PR 26-OCT-1999;	99US-0161359P.
PR 27-JUL-1999;	99US-0145919P.	PR 26-OCT-1999;	99US-0161360P.
PR 27-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146386P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161992P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161993P.
PR 03-AUG-1999;	99US-0147038P.	PR 29-OCT-1999;	99US-0162142P.
PR 04-AUG-1999;	99US-0147204P.		
PR 04-AUG-1999;	99US-0147302P.		
PR 05-AUG-1999;	99US-0147192P.		
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147493P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		
PR 13-AUG-1999;	99US-0148684P.		
PR 16-AUG-1999;	99US-0149368P.		
PR 17-AUG-1999;	99US-0149175P.		
PR 18-AUG-1999;	99US-0149426P.		
PR 20-AUG-1999;	99US-0149722P.		
PR 20-AUG-1999;	99US-0149723P.		
PR 20-AUG-1999;	99US-0149929P.		
PR 23-AUG-1999;	99US-0149902P.		
PR 23-AUG-1999;	99US-0149930P.		
PR 25-AUG-1999;	99US-0150566P.		
PR 26-AUG-1999;	99US-0150884P.		
PR 27-AUG-1999;	99US-0151065P.		
PR 27-AUG-1999;	99US-0151066P.		
PR 27-AUG-1999;	99US-0151080P.		
PR 30-AUG-1999;	99US-0151303P.		
PR 31-AUG-1999;	99US-0151438P.		
PR 01-SEP-1999;	99US-0151930P.		
PR 07-SEP-1999;	99US-0152363P.		
PR 10-SEP-1999;	99US-0153070P.		
PR 13-SEP-1999;	99US-0153758P.		
PR 15-SEP-1999;	99US-0154018P.		
PR 16-SEP-1999;	99US-0154039P.		
PR 20-SEP-1999;	99US-0154779P.		
PR 22-SEP-1999;	99US-0155139P.		
PR 23-SEP-1999;	99US-0155486P.		
PR 24-SEP-1999;	99US-0155559P.		
PR 28-SEP-1999;	99US-0156458P.		
PR 29-SEP-1999;	99US-0156596P.		
PR 04-OCT-1999;	99US-0157117P.		
PR 05-OCT-1999;	99US-0157753P.		
PR 06-OCT-1999;	99US-0157865P.		
PR 07-OCT-1999;	99US-0158029P.		
PR 08-OCT-1999;	99US-0158232P.		
PR 12-OCT-1999;	99US-0158369P.		
PR 13-OCT-1999;	99US-0159293P.		
PR 13-OCT-1999;	99US-0159294P.		
<hr/>			
Query Match 6.5%; Score 66.4; DB 3; Length 796;			
Best Local Similarity 51.3%; Pred. No. 8.5e-10;			
Matches 154; Conservative 0; Mismatches 146; Indels 0; Gaps 0;			
QY 16	TTCCATAAGTTGATTTTCTCATCCACTATCCAGAAAACGCTGTGAGGGTCCCAAGTAAG 75		
DB 113	TTCTTCAAGTTGGTCTTACCTTCAACGATGAAGATGAAGATGATGAGGATTCGCCAAGG 172		
QY 76	TTTGTGAGTAAATCAAGGATGAGCTTTGGTTCTGCTTGCACACTCAGTACCTGATGGT 135		
DB 173	TTCTGTAAGCTACAAGGGTCTAAACTCTCAGAGTTGTGTACATTAGTACCTCTCCGCGGA 232		
QY 136	CATGTTTGGCGTGTAGGACTAAGGAAAGCTGACAAACAAATTTGGTTTCAAGATGGTTGG 195		
DB 233	TATAAGCGTTCCATAAGCTTAAACGGATCGGTGAGAGATTGGTTCCATGAGGTTGG 292		
QY 196	CAAGAGTTTGTGACCGTTACTCCATTGCGATTTGGTTATCTTTTGATTTTGTAGATATGAA 255		
DB 293	AGTGAGTTTGCAGAGGCTCATTTCCATCGAGGAGGCCACTTTCTCTTTGTTGGAATACAAG 352		
QY 256	GGAACCTGCGCTTCAGGCTCTACATTTTCAATTTATCCCACTCTGAGATCAATTACCAT 315		
DB 353	AAGAATTCGAGTTTCCGTGTAATAATCTTCAATGCTTCGGCTTCGGAGACCAACTATCCT 412		
RESULT 15			
ABX62454/c			
ID ABX62454 standard; DNA; 411 BP.			
XX AC ABX62454;			
XX DT 25-FEB-2003 (first entry)			
XX DE Arabidopsis thaliana expressed sequence related polynucleotide #569.			
XX KW Transgenic plant; plant; genetically modified cell; environmental stress;			
XX KW ribozyme creation; disease resistance; stress tolerance;			
XX KW fungicide screening; insecticide screening; gene; ds.			
XX OS Unidentified.			
XX PN US2002040490-A1.			
XX PD 04-APR-2002.			
XX PF 26-JAN-2001; 2001US-00770423.			

XX 27-JAN-2000; 2000US-0178512P.  
 XX (GORL/) GORLACH J.  
 PA (ANYV/) AN Y.  
 PA (HAMI/) HAMILTON C M.  
 PA (PRIC/) PRICE J L.  
 PA (RAIN/) RAINES T M.  
 PA (YUYT/) YU Y.  
 PA (RAME/) RAMEAKA J G.  
 PA (PAGE/) PAGE A.  
 PA (WATH/) WATHEN A V.  
 PA (LEDE/) LEDFORD B L.  
 PA (WOES/) WOESSNER J P.  
 PA (HAAS/) HAAS W D.  
 PA (GARC/) GARCIA C A.  
 PA (KRIC/) KRICKER M.  
 PA (SLAT/) SLATER T.  
 PA (DAVI/) DAVIS K R.  
 PA (ALLE/) ALLEN K.  
 PA (HOFF/) HOFFMAN N.  
 PA (HURB/) HURBAN P.  
 XX  
 PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;  
 PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
 PI Hurban P;  
 XX  
 WPI; 2003-110411/10.  
 XX  
 XX Novel Arabidopsis thaliana nucleic acid useful for identifying homologous  
 PT or related genes, and to create genetically modified and transgenic  
 PT organisms, such as plant cells and plants.  
 XX  
 PS Claim 1; SEQ ID NO 569; 43pp; English.  
 XX  
 CC The invention describes an Arabidopsis thaliana nucleic acid (I). The  
 CC polypeptide (II) encoded by (I), transgenic plant (III) or genetically  
 CC modified cell (IV) are useful for screening a candidate agent for its  
 CC biological effect, by combining the candidate agent with (II), (III) or  
 CC (IV), and determining the effect of the candidate agent on (II), (III) or  
 CC (IV). (I) is useful for identifying homologous or related genes, for  
 CC producing compositions that modulate the expression or function of its  
 CC encoded protein, for mapping functional regions of the protein, in  
 CC diagnosis, for studying associated physiological pathways, for genetic  
 CC manipulation of cells, preferably plant cells, in screening assays of  
 CC various plant strains to determine the strains that are capable of  
 CC withstanding a particular disease or environmental stress, for enhancing  
 CC or inhibiting production of biosynthetic product in a plant, for  
 CC producing polypeptides, as probes for the detection of mRNA in biological  
 CC samples, to generate additional copies of (I), to generate ribozymes or  
 CC oligonucleotides, as single stranded DNA probes or as triple-strand  
 CC forming oligonucleotides, and to create genetically modified and  
 CC transgenic organisms, such as plant cells and plants. (II) or (III) is  
 CC useful for introducing or improving disease resistance and stress  
 CC tolerance in plants, screening biological active agents, e.g.,  
 CC fungicides, insecticides, etc., and for elucidating biochemical pathways.  
 CC (III) is useful as crops for their enhanced diseased resistance, enhanced  
 CC traits of interest, for screening programs, as crops which exhibit  
 CC enhanced tolerance to environmental stress, or to produce a factor. This  
 CC sequence represents a nucleic acid that may correspond to naturally  
 CC occurring Arabidopsis thaliana expressed sequences. Note: The sequence  
 CC data for this patent did not form part of the printed specification, but  
 CC was obtained in electronic format directly from the US patent office at  
 CC seqdata.uspto.gov/sequence.html?DocID=999909770423  
 XX  
 SQ Sequence 411 BP; 155 A; 92 C; 57 G; 107 T; 0 U; 0 Other;

Query Match 6.1%; Score 63; DB 8; Length 411;  
 Best Local Similarity 82.8%; Pred. No. 7e-09;  
 Matches 72; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
 QY 940 GAAGGAGAGCTGTGTGTGTTGAGCTGCTCAGAACCCAGAGATTTCGTTTGAAGTGACA 999

Db 411 GAAGGAGAGTGTGTGTGTTGAGCTGCTCAGAACCTCGGGATTTCGTTCTCGAAGTCACC 352  
 QY 1000 GCCTTTTCGAGTCAACGAGTACGCTCTGA 1026  
 Db 351 GCCTTTTCGAGTCAATGAGTATGTGTGA 325  
 Search completed: December 30, 2004, 13:26:53  
 Job time : 605 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 30, 2004, 12:43:00 ; Search time 113 Seconds  
(without alignments)  
6460.002 Million cell updates/sec

Title: US-10-088-187A-10\_COPY\_269\_1295  
Perfect score: 1027  
Sequence: 1 atgccacgcctttcttcca.....agtcacgagtagctgtgaa 1027

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2.6/prodata/1/ina/5A COMB.seq:  
2: /cgn2.6/prodata/1/ina/5B COMB.seq:  
3: /cgn2.6/prodata/1/ina/6A COMB.seq:  
4: /cgn2.6/prodata/1/ina/6B COMB.seq:  
5: /cgn2.6/prodata/1/ina/6C COMB.seq:  
6: /cgn2.6/prodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	43.2	4.2	7218	1	US-08-232-463-14
C 2	43	4.2	832	4	US-09-621-976-2813
C 3	40.2	3.9	1141	4	US-09-806-708B-22
C 4	39.2	3.8	7218	1	US-08-232-463-14
C 5	35.8	3.5	4143	4	US-09-328-352-4006
C 6	35.8	3.5	4488	4	US-08-956-171B-228
C 7	35.8	3.5	4488	4	US-08-781-986A-228
C 8	35.2	3.4	3441	3	US-08-896-164-85
C 9	34.6	3.4	832	4	US-09-621-976-2813
C 10	34.6	3.4	1664976	4	US-08-916-421B-1
C 11	34.6	3.4	1664976	4	US-09-692-570-1
C 12	34.4	3.3	1494	4	US-09-107-532A-2145
C 13	34.2	3.3	710	4	US-09-270-767-29582
C 14	34.2	3.3	1421	4	US-09-270-767-13580
C 15	34.2	3.3	3492	3	US-08-923-992A-9
C 16	34.2	3.3	6908	4	US-09-620-312D-977
C 17	34	3.3	3025	4	US-08-961-527-168
C 18	34	3.3	18431	4	US-09-221-017B-1090
C 19	33.6	3.3	499	4	US-09-270-767-358
C 20	33.6	3.3	499	4	US-09-270-767-15640
C 21	33.6	3.3	2535	4	US-09-799-451-549
C 22	33.4	3.3	314	4	US-09-614-912-151
C 23	33.4	3.3	1734	4	US-09-248-796A-6833
C 24	33.2	3.2	152331	3	US-09-128-155-16
C 25	33.2	3.2	176373	3	US-09-128-155-17
C 26	32.8	3.2	418	4	US-09-270-767-11173
C 27	32.8	3.2	2061	4	US-09-107-532A-1348

28	32.6	3.2	2617	1	US-08-430-024-1	Sequence 1, Appl
29	32.6	3.2	2617	1	US-08-782-009-1	Sequence 1, Appl
30	32.6	3.2	2617	3	US-09-017-302-1	Sequence 1, Appl
31	32.6	3.2	3294	3	US-08-923-992A-7	Sequence 7, Appl
32	32.6	3.2	3384	3	US-08-923-992A-5	Sequence 5, Appl
33	32.6	3.2	4200	1	US-08-242-932-1	Sequence 1, Appl
34	32.6	3.2	4200	1	US-08-714-481-1	Sequence 1, Appl
35	32.6	3.2	4200	3	US-08-923-992A-1	Sequence 1, Appl
36	32.6	3.2	4200	5	PCT-US95-06111-1	Sequence 1, Appl
C 37	32.6	3.2	90050	3	US-09-245-041-5	Sequence 5, Appl
C 38	32.6	3.2	90050	4	US-09-358-058B-5	Sequence 5, Appl
C 39	32.6	3.2	90050	4	US-09-893-238-5	Sequence 5, Appl
40	32.4	3.2	485	3	US-09-177-325-4	Sequence 4, Appl
41	32.4	3.2	485	3	US-09-411-812A-4	Sequence 4, Appl
42	32.4	3.2	485	3	US-09-590-113-4	Sequence 4, Appl
C 43	32.4	3.2	1141	4	US-09-806-708B-22	Sequence 22, Appl
C 44	32.4	3.2	2922	2	US-08-683-262B-74	Sequence 74, Appl
C 45	32.4	3.2	2922	3	US-09-361-707-74	Sequence 74, Appl

## ALIGNMENTS

RESULT 1  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Iardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgt-F1s  
US-08-232-463-14

Query Match 4.2%; Score 43.2; DB 1; Length 7218;



```
;
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
;
US-08-232-463-14

Query Match 3.8%; Score 39.2; DB 1; Length 7218;
Best Local Similarity 1.7%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 5; Conservative 176; Mismatches 119;

QY 202 TTGTGACCGTTACTCCATTCGATGTTATCTTTGATTTAGATGAAGAAAC 261
DB 1141 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1200

QY 262 TCTGCTTCAGCGCTGACATTTTCAATTTATCCACTGTGAGATCAATACCACCC 321
DB 1201 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1260

QY 322 GGTCTCATGATTCGCTCACACCACTTCAACGCGCGGTTGTTGAAGACCTTGA 381
DB 1261 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1320

QY 382 GATGAAGATGCCGAGGTCTCTCTCTGTGTACCATCACCCTCTCTGAGTCT 441
DB 1321 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1380

QY 442 ACAGTACCAACCAAGAGGTATGCTAGTTCAGCATCAACACTTGTCTACTGACCA 501
DB 1381 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1440

RESULT 5
US-09-328-352-4006/c
; Sequence 4006, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

;
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4005
; LENGTH: 4143
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
;
US-09-328-352-4006

Query Match 3.5%; Score 35.8; DB 4; Length 4143;
Best Local Similarity 52.3%; Pred. No. 1.1; Indels 0; Gaps 0;
Matches 79; Conservative 0; Mismatches 72;

QY 197 AAGAGTTTGTGACCGTTACTCCATTCGCAATTCGATTCGTTATCTTTGATTTAGATGAAG 256
DB 2342 ATGAGTTATAGATATCTACACCTGCCTCACCTGCAACCAATTCATCTTCGTTTACACGA 2283

QY 257 GAAACTCTGCTTCAGCGTCTACATTTTCAATTTATCCCACTCTGAGATCAATTACCATT 316
DB 2282 ATAAGCATACGAGAAGCATCTACATATTCATATACACCGCCACGGTTTCGATTACACAC 2223

QY 317 CCACCGGTCTCATGATTCGCTCACAACCA 347
DB 2222 ACACGAGATCAGTGCAACGTCGTTCCCA 2192

RESULT 6
US-08-956-171E-228
; Sequence 228, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248PI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 228:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4488 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
```

```

US-08-896-164-85
; Sequence 85, Application US/08896164
; Patent No. 6218521
; GENERAL INFORMATION:
; APPLICANT: OBAITA, Yuichi
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
; WITH GASTRIC CANCER AND METHODS FOR
; TITLE OF INVENTION: DIAGNOSING AND TREATING GASTRIC CANCER
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,164
; FILING DATE: July 17, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6218521man D. Hanson
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3441 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-896-164-85

Query Match          3.4%; Score 35.2; DB 3; Length 3441;
Best Local Similarity 51.2%; Pred. No. 1.5;
Matches      82; Conservative    0; Mismatches   78; Indels     0; Gaps      0;

QY       658 AGAACCGTGACTGCAGAGAAGAGAGAGAGAGAGCCATCAATGACGCCAAACGGTTCGRACCA 717
Db        1964 ATACCCCTTCTCTCCACAAAAAGAGACTGAGAGGGGAGATCAAAGTGAAGGGTCCAAGCGA 2023

QY       718 ACAAAACCCTTCTTCAGAGTGTTCTGCACCATTCTATCTATACAGAGGTTGCATCATG 777
Db        2024 ACTTAGTGACTCTTCAGGTGTTTGTGAGTTTGGCTTTTTTCTTCTTTGTGTAATCTT 2083

QY       778 TAFTCTTCCTTCTGGGTTTCTCGAAGAGTCACTTAAGTGGGA 817
Db        2084 TAGTGATTGCTTGATGTACTTAATAATTACCTGAGTTGA 2123


RESULT 9
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S. J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match      3.4%; Score 34.6; DB 4; Length 832;
Best Local Similarity 9.6%; Pred. No. 0.99;
Matches 30; Conservative 153; Mismatches 127; Indels 1; Gaps 1;

QY 331 GATTCGCTCACACACACATTCACACGCGCGTGTGTTGACAGACCTTGAAGATGAGAT 390
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18 KWKWSWSYWMYKWTYKTYWRWRKKAAMKWKYKWTWYRYAMWGTGTYKKACMCRPK 77
QY 391 GCGAGGTCTATCTTCTCTCTGTCACCATCACCACTCTCTGAGTCTACAGTACCA 450
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
78 TKKKKKKGYMMYWGWSYMAWTRTWGAYYSMMYWRRCWKKKAYRKTTCYS 137
QY 451 GCCAACAAAGGGTATGCTAGTTCAGCCATCCAAACCTGTTCACTGGACAGTTAAAGCT 510
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
138 SKGWTWKKKKAATTTWKKTY-YWAATRYMMCMCTKKRWRASWYCWNGKARKWSTW 196
QY 511 GAAGACCCACGCCAACCCCAAAATACCTAAAGAGAGAGGGAGGAGAGAAAATGCT 570
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
197 RKSRSYASARSACRCCYSCSGAMSKYMWWRMGWATGAGMKAWRASCMWRRKYAGK 256
QY 571 GATCTCTGAGGAATAAATCATCATGCTCCGCGAGATGATCCAGAGAACCGTTCAAG 630
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
257 SKTSYKSMWCMWTRSKYCYTKARWTGYCYRKGGMWGRGRWYASKYKTYMWRWNCWAR 316
QY 631 TTCTACGAGAG 641
Db   : : : : :
317 MYRYSRGTAS 327

RESULT 10
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28232)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)

; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
```

Query Match 3.4%; Score 34.6; DB 4; Length 1664976;  
Best Local Similarity 48.3%; Pred. No. 80;  
Matches 97; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

LOCATION: (98266)..(98266)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (98343)..(98343)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (103998)..(103998)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (148948)..(148948)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (163385)..(163385)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (191989)..(191989)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (191995)..(191995)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (231980)..(231980)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (234187)..(234187)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (234220)..(234220)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (234814)..(234814)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (309398)..(309398)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (309418)..(309418)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (312837)..(312837)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (312993)..(312993)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (319226)..(319226)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (559167)..(559167)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (559241)..(559241)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (600992)..(600992)

OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (622708)..(622708)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (682442)..(682442)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (713652)..(713652)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (741684)..(741684)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (779455)..(779455)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (779676)..(779676)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (855539)..(855539)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (871619)..(871619)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1119881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g

US-09-107-532A-2145

Query Match 3.4%; Score 34.6; DB 4; Length 1664976;  
Best Local Similarity 48.3%; Pred. No. 80;  
Matches 97; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 506 AAGCTGAAGAGCGCAACGCCAACCCCAAAATACCTTAAAGAGAGGGAGGAAGAAAA 565  
DB 547953 AAGTTGGAGATTATCGGATATAAGAAATAAATTTTGGAGAGAGCAATACAAATTTAA 547894

QY 566 ATGCTGATCCTGAGGAAATAAATCACTCAGCTCCCGGAGATGATCCAGAGAACCGTT 625  
DB 547893 AAGATGGTAAATAGAGAAATCTCATCAGGTTTGGAAATGGTGTAGCTGTAGAGTCT 547834

QY 626 CAAAGTTCTACGAGAGTGCTTCTGGCAGAAAGAACCGTGACTCCAGAGAAAGAGAGA 685  
DB 547833 TATACAAAATGGATGGGGTTTGTATACATCAACATAGTGAGTGAAGAGGAATCGAAA 547774

QY 686 GAGCCATCAATGAGCGCAAAA 706  
DB 547773 AACTCATAAATAAGCGGTATA 547753

RESULT 12

US-09-107-532A-2145/c  
; Sequence 2145, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD/ROM ISO9660  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 2145:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1494 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (B) LOCATION 1...1494  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2145:

US-09-107-532A-2145

Query Match 3.3%; Score 34.4; DB 4; Length 1494;  
Best Local Similarity 59.0%; Pred. No. 1.6;  
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 573 TCCTGAGGAAATAAATCACTCAGCTCCGCGAGATGATGATCCAGAGAACCGTTCAAAGTT 632  
DB 513 TCCTGATGAAATCGCAGCAGAGCTCTCTCACTTGATCCGCCCGGTGTACATTTCAAGGTT 454

QY 633 CTACGAGAGTGCTTCTTCGAGAAAGAGAACCGTGACTGCA 672  
DB 453 CCAAGGGTTTCTTGCTGGGCCATAGATGACTGGGTCTGTA 414

RESULT 13

US-09-270-767-29582  
; Sequence 29582, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 29582  
; LENGTH: 710  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
; US-09-270-767-29582

Query Match 3.3%; Score 34.2; DB 4; Length 710;  
Best Local Similarity 45.9%; Pred. No. 1.2;  
Matches 117; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

QY 389 ATGCCGAGGTCATCTTTCTTCTTCTGTGTACCCATCACCCTTCCTGAGTCTACAGTAC 448  
DB 78 ATACAAAAACAGCCCAATCTCTGCTGTAAAAATTTGCCAATTTATAAACAAAAATGATC 137

QY 449 CAGCCAAACAAAGGTATGCTAGTTTCAGCCATCCAAACCTTTTCCTGAGGACAGTTAAAG 508  
DB 138 CAAATCAAAATAATATGCAAAACACTAACTGTTTAAAGCAGGAGCTGGAAATGAAGTTG 197

QY 509 CTGAAGAGCCCAACGCCAACCCCAAAATACCTTAAAAAGAGAGGAGGAGGAGAAAAATG 568  
DB 198 CAAAAACACACGCCGAGCAAAATAAAAACCTGCAATGGCGGTAAAAACGAAAAAGCACTG 257

QY 569 CTGATCTCTGAGAAATAAATCACTCAGCTCCGCGAGATGATGATCCAGAGAACCGTTCAA 628  
DB 258 CTACCCCTAAGGTAGTAAATAATCCAAACACTCAAAAAAATATTTCAGAAATAGCAACAA 317

QY 629 AGTTCTACGAGAGTG 643  
DB 318 ACAITCAAAAGAAATG 332

RESULT 14

US-09-270-767-13580  
; Sequence 13580, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 13580  
; LENGTH: 1421  
; TYPE: DNA

ORGANISM: Drosophila melanogaster  
US-09-270-767-13580

Query Match 3.3%; Score 34.2; DB 4; Length 1421;  
Best Local Similarity 45.9%; Pred. No. 1.8;  
Matches 117; Conservative 0; Mismatches 138; Indels 0; Gaps 0;  
QY 389 ATGCGAGGTGATCTTCTTCTTCTGTGTGATCCATCACCCTTCTGTGATCTTACAGTAC 448  
DB 789 ATAACAAAAACAGCCCAATCTCTGCTGTGTTAAAAATTTGCCAATTATAAACAAAAATGTAC 848  
QY 449 CAGCCAAACAAAGGTATGTTGATGTTGATGATCCATCACAACCTTGTTCATCTGGACCAAGTTAAAG 508  
DB 849 CAAATCAAAATAATAATGCAACACATTAACGTTGTTAAGCAGCGAGCTGGAAATGAAGTTG 908  
QY 509 CTGAAGAGCCCAACGCCCAACCCCAAAATACCTAAAAAGAGAGGGAGGAGAAAGAAAAATG 568  
DB 909 CAAAACACAGCCGACGCAAAATATAAAACCTGCAATGGCGGTAAACGAAAGCACTG 968  
QY 569 CTGATCCTGAGGAAATAAACTCATCTCAGCTCCGAGATGATGATCCAGAGAACCGTTCAA 628  
DB 969 CTACCCCTAAGGTAGTAATAATCAACACTCAAAAAAATATTTCAGATAGCACAAACA 1028  
QY 629 AGTTCTACGAGTG 643  
DB 1029 ACATTCANAAGAATG 1043

RESULT 15  
US-08-923-992A-9  
Sequence 9, Application US/08923992A  
Patent No. 6280738  
GENERAL INFORMATION:  
APPLICANT: Tai, Joseph Y.  
APPLICANT: Blake, Milan S.  
TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B  
TITLE OF INVENTION: Streptococcal Beta Antigens  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/923,992A  
FILING DATE: 05-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/024,707  
FILING DATE: 06-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 1438.0140001/RWE  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3492 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS

LOCATION: 1..3492  
US-08-923-992A-9  
Query Match 3.3%; Score 34.2; DB 3; Length 3492;  
Best Local Similarity 53.3%; Pred. No. 3.1;  
Matches 72; Conservative 0; Mismatches 63; Indels 0; Gaps 0;  
QY 457 AAAGGGTATGCTAGTTTCAGCCATCCAAACCTTTGTTTCACTGGACCACTTAAAGCTGAAGAG 516  
DB 1552 AAAGATGATTTCTGGCAATGTAGTTCAAAAAACATTTCACCTATTACAGTGCAGAAAGAGAG 1611  
QY 517 CCAAGCCCAACCCCAAAATACCTTAAAGAGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAG 576  
DB 1612 GAGAAACCAAGTTTCTTAAACACCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1671  
QY 577 GAGGAAATAAACTCA 591  
DB 1672 CAAGAACCAAAATCA 1686  
Search completed: December 30, 2004, 14:34:03  
Job time : 116 secs

**THIS PAGE BLANK (USPTO)**



Db 322 ACCGTTCTGATGGTCAATGTTTGGCGGTGAGGAATAAAGAAAGTTGACAAACAGGCTTGG 381  
Qy 181 TTTCAAGATGGTGGCAAGAGTTTCTTACCGTTACTCCATTCGCAATGGTTATCTTTTG 240  
Db 382 TTTAAGGAGGGTGGCAGGAATTTTGCAGAACCTTACTATATCCGTTGGCTACTTTTTG 441  
Qy 241 ATTTTATAGATATGAAGAACTCTGCTTCCAGCGTCTACATTTTCAATTTTATCCACTCT 300  
Db 442 ATTTTCAGATATGAAGGAATTTCCGTTTCAGTGTAGCATATTTGATTTGTACAACTCT 501  
Qy 301 GAGATCAATTAACATTTCCACCGGCTCTATGATTCGCTCAACACCACTTCAACCGGCC 360  
Db 502 GAAATAAATCATCAACAAATCTCTTGTGTACTCAATACAATCTCGGAAGACAATAT 561  
Qy 361 CGTTTGTGTTGAAGACTTGAAGATGAAGATGCCGAGGTCACTTTCCTTCTCTGT---- 416  
Db 562 CCATTTGAGAACTTTGAAGATGATGAATGTGTCTCTCCAGCGCTGCCTAAATTTGTTGGT 621  
Qy 417 -----GTACCCTATCACCACCTTCCTGAGT 439  
Db 622 GGGTCAAACTTAATCTGATATAAATCTGGAGTGGAGAGTCAACCATCAGCACCCTAAGGT 681  
Qy 440 CTACAGTACACCCCAACAAAGGATGCTAGTTTCAAGC-----ATCCAAACCTTG 489  
Db 682 GTTAATAATCAACCTATTTCGGGGTATGGCTTTTCACTTACTGTATTTGAATTTAAAGAT 741  
Qy 490 TTCACCTGACGATTAAGCTGAAGCAAGCAAGCCCAACCCCAAAATACCTTAAAGAGA 549  
Db 742 TCTGTAATGAAGTGAATGGCAATTTCTTCAGGTGAGAGCTACCAAAACTGAAAAACCA 801  
Qy 550 GGGAGGAAGAAATAATCTCATCTGAGGAAATAAATCACTACAGCTCCGCGAGATGAT 609  
Db 802 GGGAGGAAGAAAGCAGAGTTTGAGCTAGTGAAGAGGATTCATCTCTGGACATGAGAT 861  
Qy 610 GATCAGAGAACCGTTCAAGTTCTACGAGAGTGTCTTCGAGAGAAAGAACCGTGACT 669  
Db 862 GACATGCAATGCGTAATAGATTTTATGAAAGTGTCTCCGCCAGGAAGAAATTTGACT 921  
Qy 670 GCAGAGAAAGAGAGAGCCATCAATGCAGCAAAAGCTTCGAACCAACCAACCCCTTC 729  
Db 922 GCTGAAGAAAGAGAGAGCGATTAATGCAGCAAAAGCATTTTGAGCCTACTAACCCCTTC 981  
Qy 730 TFCAGAGTGTCTCGACCATCTCTATATACAGAGTTGATCATGTATCTCTCTCT 789  
Db 982 TCCAGGTCCTCTGGACCGTCTTATCTATACAGGGATGATTTATGTACTTACCATCG 1041  
Qy 790 GGGTTGCTGAGAGTACCTAAGTGGGATCTCCGGGTTTCATCAAAAGTCCAGCTTCGGA- 848  
Db 1042 TGCTTTGCTGAGCAACATCTAAGCGGGTTTCAGGATTCATTAACCTTCAGCTCCGGAT 1101  
Qy 849 --GAAACAAATGGCCTGTGATGTCTCTACAAAGCCGGAGAGCCAAATTCAGTCAAGA 906  
Db 1102 GGTAGACACGCGCCTGTTTCGATGCGGTTATAGAGGAGGAGGCTTAAGTTTCAGTCAAGA 1161  
Qy 907 TGGTACGAATTCACCTCAGAGAACTTTAGAGAGAGAGAGCTGTGTGTGTGTGAGCTG 966  
Db 1162 TGGTATGAATTTACATTTGGAGAACAAATTTGGGGAGAGAGATGTCTGTCTTTGAACTG 1221  
Qy 967 CTCAGAACAGAGATTTGTTTGAAGTGAAGAGCTTTTCAGTCAACGA 1016  
Db 1222 CTCAGATCGAGGAATTTGTGCTCAAGTTCACCGTATTTTCGTGTAATGA 1271

RESULT 2  
US-09-938-842A-273  
; Sequence 273, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 273  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-273

Query Match 38.7%; Score 397.6; DB 9; Length 681;  
Best Local Similarity 76.6%; Pred. No. 1.2e-115;  
Matches 538; Conservative 0; Mismatches 134; Indels 30; Gaps 3;

Qy 328 ATGATTCGCTGCAACCACTT---CAACGCGCCGCTTTTGTGAAGACCTTGAAGAT 384  
Db 7 ATGATTCGCTGCAACCACTT---CAACGCGCCGCTTTTGTGAAGATCCTGAATC 66  
Qy 385 GAAGATGCGGAGGTCACTTCTTCTCTGTGTACCCATCACCACCTTCTGAGTCTACA 444  
Db 67 AAGATGCTAAGGTCACTTATCCATCGAAC-----CCTGAATCTACT 108  
Qy 445 GTACAGCCCAACAAAGGATGCTTAGTTTCAGCCATCCAAACCTTTGTTCACTGGACAGTT 504  
Db 109 GAACCGATGAATAAGGTTATGGCGTTTACAGCCATCCAAAGCTTTTCAAGAACTCT 168  
Qy 505 AAAGCTGAAGAGCCAAACCCCAACCAAAATCTTAAAGAGAGAGGAGAGAGAGAAA 564  
Db 169 AAAGCTGAAGAG-----AACGCCCAAGGTACTTAAAGAGAGAGAGAGAGAGAG 219  
Qy 565 AATGCTGATCTGAGGAATAAATCACTACGCTCCGCGAGATGATATCCAGAGAACCGT 624  
Db 220 AATCTAATCCGAGAGAGTAACTCTTCACTCCGCTGGAGATGATCTCAGAGAACCGC 279  
Qy 625 TCAAGTTCTACGAGAGTCTTCTCGAGAAAGAGAACCGTACCTGCGAGAGAGAGAGAG 684  
Db 280 TCAAGTTCTACGAGAGTCTTCTGCTAGAAAGAGAACTGTAACTGCGAGAGAGAGAG 339  
Qy 685 AGAGCCATCAATGAGCCCAAAACGTTGCAACCAACAAACCTTTCTTCAAGAGTGGTTCTG 744  
Db 340 AGAGCCGTCATGAGCCCAAAACATTCGAACCAACAAATCTTACTTTAGAGTTGTTCTG 399  
Qy 745 CGACATCCTATCTATACAGAGGTTGATCATGTATCTTCTCTGCGGTTTGTGAGAGAG 804  
Db 400 CGACATCCTATCTATACAGAGGTTGATCATGTATCTTCTCTGCGGTTTGTGAGAGAG 459  
Qy 805 TACCTAAGTGGGATCTCCGGGTTCACTCAAAAGTTCAGCTTTCGCGAGAGAAACAATGGCCTGT 864  
Db 460 TACCTAAGTGGGATCTCTGGTTTCACTCAAGCTCCAGCTCGGTGAGAGAAACAATGGCCTGT 519  
Qy 865 CGATGTCTTCAAAAGCCGGAGAGCCAAATTCAGTCAAGATGATGATGATGATGATGATGAT 924  
Db 520 AGGTGCTCTTCAAAAGCCGGAGAGAGTAAAGTTTAGCCAAAGATGATGATGATGATGATGAT 579  
Qy 925 GAGAACACTTAGGAGAGAGAGAGTCTGTGTGTGTGAGCTGCTCAGAACCCAGAGATTTTC 984  
Db 580 GAGAACAAATATAGCGAGAGAGAGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 639  
Qy 985 GTTTTGAAGTGAAGAGAGAGAGTTCAGATCAACGAGTACGTTCTGA 1026  
Db 640 GTTCTCGAGAGTCAACCGCTTTTCGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 681

RESULT 3  
US-09-938-842A-273



## RESULT 5

US-10-767-795-1316  
; Sequence 1316, Application US/10767795  
; Publication No. US20040181830A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Zhou, Yihua  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53534)B  
; CURRENT APPLICATION NUMBER: US/10/767,795  
; NUMBER OF SEQ ID NOS: 2004-01-30  
; SEQ ID NO 1316  
; LENGTH: 1111  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE:  
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C1540\_3  
US-10-767-795-1316

Query Match 28.0%; Score 287.8; DB 17; Length 1111;  
Best Local Similarity 61.3%; Pred. No. 1.9e-80;  
Matches 573; Conservative 0; Mismatches 267; Indels 95; Gaps 3;

QY 1 ATGCCACGCCCTTTCTTCATAGTTGATTTCTCATCCACTATCCAGAAAACGTCG 60  
DB |||||||  
QY 269 ATGCCACGCCCTTTCTTCATAGTTGATTTCTCATCCACTATCCAGAAAACGTCG 328  
DB |||||||  
QY 61 AGGGTCCCAAGATTTGTGAGTAAATCAAGATGAGCTTTCGGTGTCTGTCACCTC 120  
DB |||||||  
QY 329 AGGATCCCGATAACTTCGTAAAGAAATTCGGGGACGAACTTCTGTCGCGCTCTC 388  
DB |||||||  
QY 121 ACAGTACCTGATGGTCATGTTGGCGTGTAGACTAAGAAAGCTGACAAACAAATTTGG 180  
DB |||||||  
QY 389 ACTGTTCTCTGACGGTCATGTTTGGCGTGTAGAAATAAGAAAGCGACCAACAAAGTTGG 448  
DB |||||||  
QY 181 TTTCAAGATGTTGGCAAGATTTGTGACCGTTACTCCATTCGCATCGCATGTTGTTTGTG 240  
DB |||||||  
QY 449 TTTACAGAGGTTGGCAGAGTTCTAGAGCGGTACCATATCCGTTGTTGGCTACGTACTG 508  
DB |||||||  
QY 241 ATTTTATAGATATGAAGAACTCTGCTTTACGCGTCTACATTTTCAATTTTATCCACTCT 300  
DB |||||||  
QY 509 GTTTTCAGATATGAATGGAATCTTGTGTTTCAATAAAGCGTATCCATTGA ----- 559  
QY 301 GAGATCAATTACCATTCACCGGCTCTATGGAATTCGGTCTCAACCACTTCAACCGCGC 360  
DB ----- 559  
QY 361 CGTTGTTTGAAGACCTTGAAGATGAAGATCGGAGGTGATCTTTCTCTCTCTGTATAC 420  
DB |||||||  
QY 560 -----AGACTTGAAGATGACGAATGATGATCTCTCCGGCACTGCGAAT 602  
DB |||||||  
QY 421 CCATCACCACCTTCTGAGTCTACAGTACCAGCCCAACAAAGGTATGCTAGTTTCAGCCATC 480  
DB |||||||  
QY 603 TTGT-----TTTCTGCGTCTAAAGTTAATACTGATCACTGAGTGTGAATCAAT 656  
DB |||||||  
QY 481 CAAACCTGTTTCACTGGACGATTAAGCTGAAGAGCCCAACGCCCAACCCCAAAATACCT 540  
DB |||||||  
QY 657 TTTCTGCGCTTTAAAGGCGATGAATAGCCAACTCTTTTCGAGGTGCGATGCTACCAAAACCA 716  
DB |||||||  
QY 541 AAAAGAGAGCGGAGAGAAAGAAATGCTGATCTCTGAGGAATAAATCACTCATCAGCTCG 600  
DB |||||||  
QY 717 AACAGCCTGGGAGAGAAAGCAGAGTTTGTATCAGCTGAACCGGATTCATCCGTTGGA 776  
DB |||||||  
QY 601 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGTGTCTTCTGCGAGAAAGAGA 660  
DB |||||||  
QY 777 CGGAAATATGATGTGTATACAACTTTAGATTATACGAAGTGTCTTACGCCGGAAGAGA 836  
DB |||||||  
QY 661 ACCGTGACTGCAAGAAAGAGAGAGAGCCATCAATGACGCCAAACCGTTGCAACCAACA 720  
DB |||||||  
QY 837 ACCGTGACAAACGAAAGAGAGAGAGGCGCAATTAACGCGAGCCAAATCGTTTGAGCGGATG 896  
DB |||||||

## RESULT 6

US-10-767-795-1317  
; Sequence 1317, Application US/10767795  
; Publication No. US20040181830A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Zhou, Yihua  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53534)B  
; CURRENT APPLICATION NUMBER: US/10/767,795  
; NUMBER OF SEQ ID NOS: 2004-01-30  
; SEQ ID NO 1317  
; LENGTH: 1062  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE:  
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C1540\_2  
US-10-767-795-1317

Query Match 24.8%; Score 255.2; DB 17; Length 1062;  
Best Local Similarity 73.0%; Pred. No. 4.6e-70;  
Matches 356; Conservative 0; Mismatches 128; Indels 4; Gaps 2;

QY 543 AAAGAGAGGAGGAGAGAAATGCTGATCTCTGAGGAATAAATCACTCATCAGCTCCGCG 602  
DB |||||||  
QY 164 AAAAAGCGTGGGAGGAGCGGAAGTTTGAATCTTAACTGCGAGGATTCATCTGCTGACG 223  
DB |||||||  
QY 603 AGATGATGATCCAGAGAACCGTTTCAAAGTTCTACGAGAGTCTTCTGCGAGAAAGAGAAC 662  
DB |||||||  
QY 224 TGAAGATGATGCTGACATGCGCTTAGTGTCTATGAAAGTCTTCTGCGGAAAGAGAAC 283  
DB |||||||  
QY 663 CGTGAATGCAAGAAAGAGAGAGAGCCATCAATGAGCCCAAAAGCTTTCGAAACCAACAAA 722  
DB |||||||  
QY 284 CGTGACAGCTGAAGAAAGAGAGAGAGCAATCAATGCGGCCAAAGCAATTTGAGCCCTACTAA 343  
DB |||||||  
QY 723 CCCTTTCTCAGAGTGGTTCTGCGACCATCTATCTATACAGAGTTCATCATGTATCT 782  
DB |||||||  
QY 344 CCCTTTCTGCGAGGTCGCTTTGCGACCATCATATCTGTACAGGGGATGCAATATGTACTT 403  
DB |||||||  
QY 783 TCCTTTCTGGGTTTCTGAGAAAGTACCTAAAGTGGGATCTCCGGGTT-CATCAAAGTCCAGC 841  
DB |||||||  
QY 404 ACCGTGCTGCTTCTGAGAGAGCATCTAAGTGGGTTTCTGATTTCCATTAACCTTCAGC 463  
DB |||||||  
QY 842 TTGCGGA---GAAACAAATGGCTGTTTGCATGTCTCTACAAAGCCGGAGAGAGCCAAATTC 898  
DB |||||||  
QY 464 TTCTGTATGGGAGACAGTGGTCTGTACGATCTCGTTATAAAGAGGCGCAAGCTTAAGTTCA 523  
DB |||||||  
QY 899 GTCAAGGATGCTACGAATTCATCTAGAGAACACTTAGGAGAGAGAGACCTCTGTGTGT 958  
DB |||||||  
QY 524 GTCAAGGATGCTACGAATTTACATTTGAGAGAAATAATTTGGGAGAGAGGATGTCTGTGTCT 583  
DB |||||||  
QY 959 TTGAGCTGCTCAGAAACAGAGATTTTCTGTTTGAAGATGACAGCCCTTTTCGAGTCAACGAGT 1018  
DB |||||||

Db 584 TCAGCTGCTCAGATCGAGGAATTCGTGCTCAAAAGTTACCGTATTTCTGTTGAAGGAAA 643  
QY 1019 ACCTCTGA 1026  
Db 644 GCCTGGA 651

## RESULT 7

US-10-424-599-118105  
; Sequence 118105, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 118105  
; LENGTH: 2141  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_77658C.1  
US-10-424-599-118105

Query Match 22.1%; Score 227.2; DB 16; Length 2141;  
Best Local Similarity 71.4%; Pred. No. 6.2e-61;  
Matches 314; Conservative 0; Mismatches 123; Indels 3; Gaps 1;  
QY 590 CATCAGCTCGGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGAGTGTCTTG 649  
Db 1203 CCTCTCTCGGCACGAGAGGAGGTAGAAATGCGCTTTAGATTTTATGAAGTGTCTTG 1262  
QY 650 CGAGAAAGAGAACCGTGACTGCAGAGAAAGAGAGAGCCATCAATGACGCCAAACGT 709  
Db 1263 CAGAAAGAAAGACTGTGACAGCTGAAGAAAGAGAGAGGTGTCATCAAGCAAGCAAAAGCAT 1322  
QY 710 TCGAACCAACAAACCCCTTTCTTCAGAGTGTCTGCGACCATCTCTATATACAGAGTT 769  
Db 1323 TTGAACCATTAATCTTCTCTGAGTGTCTCTGCGCCCTCATATTTATATAGGGAT 1382  
QY 770 GCATCATGTATCTCTCTCTGAGTGTCTGAGAGTACCTAAGTGGGATCTCCGGGTTCA 829  
Db 1383 GCATATATGTATCTGCCATCATGCTTTGCAGAGAGCATTTGAATGGAGTTTCAGGGTTTA 1442  
QY 830 TCAAGTCCAGCT---TGGGAGAAACATGGCTGTTTCGATGTTCTCTACAAAGCCGGA 886  
Db 1443 TTAACCTTCAGATCTCTAATGTTAGACAGTGGCCGGTTCGCTGCCCTTTATTAAGGAGGTA 1502  
QY 887 GAGCCAAATTCAGTCAAGGATGTTAGCAATTCATCTAGAGAACAACTTAGGAGAGGAG 946  
Db 1503 GAGCCAAAGTTAAGCAGGGGTGTTGAATTTTCATTAGAGAACAACTTAGGAGAGGTTG 1562  
QY 947 AGCTCTGTGTGTGAGTGTCTCGTCAAGAACAGAGATTTCTGTTTGAAGTGCAGCGCTTTC 1006  
Db 1563 ATGCTGTGTGTGTGAGTGTCTCGTCAAGAACAGAGTGTGCTGCAAGTTACCATCTTTC 1622  
QY 1007 GAGTCAACGAGTACGTTCA 1026  
Db 1623 ATGTTACCGAGGATGGGA 1642

## RESULT 8

US-10-424-599-97592  
; Sequence 97592, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 97592  
; LENGTH: 1760  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_59139C.1  
US-10-424-599-97592

Query Match 22.0%; Score 226.2; DB 16; Length 1760;  
Best Local Similarity 72.6%; Pred. No. 1.1e-60;  
Matches 307; Conservative 0; Mismatches 113; Indels 3; Gaps 1;  
QY 602 GAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGAGTGTCTTTCGCGAAGAGAA 661  
Db 946 GTGAANAATGTGCCGATTAAGCCCTGGAGGTTTTATGAACTGCGCATCTGCAAGGAAAAGTA 1005  
QY 662 CGTGACTGCAGAGAAAGAGAGAGAGCCATCAATGCGCCAAAAGTTTCGAACCAACAA 721  
Db 1006 CCGTGACAGCAGAGAAAGAGAGAGAGGCAATTAATGATCAAAAACATTTGAACCAACTA 1065  
QY 722 ACCCTTTCTCAGAGTGGTTCTGCGACCATCTATCTATACAGAGTTGTCATCATGTATC 781  
Db 1066 ATCTTTCTCGGAGTGTGCTCGACCCCTCTATTTGATAGGGATGTCATAATGTATC 1125  
QY 782 TTCCTTCTGGGTTTCTGAGAGTACCTAAGTGGGATCTCCGGGTTTCATCAAGTCCAGC 841  
Db 1126 TSCCTTCTGCTGTTCTGAAAAGAAATTTGAATGGGGTTTCGGGATTCATCAAACTTCAGT 1185  
QY 842 TTGC---GGAGAAACAATGGCTGTTTCGATGTTCTACAAAGCCGGAGAGCCAAATTC 898  
Db 1186 TGTCAACGGTTAGACAGTGTCTGTTGCTGCTCTATAGGGAGGTGCGAGCCAAAGTTAA 1245  
QY 899 GTCAAGGATGTGCAAAATTCATCTAGAGAACAACTTAGGAGAGAGAGCGTCTGTGTGT 958  
Db 1246 GCCAAGGTTGGTTTGAATTCACAGTGGAGAACAAATTTGGGAGAGGTGACGCTGTGTGT 1305  
QY 959 TTGACTGCTCAGAACCAAGAGATTTTCGTTTGAAGTGCAGCCCTTTTCGAGTCAAGGAGT 1018  
Db 1306 TTGAGCTCTTAGAACGAAGGAGTGTGTCTGCAAGTTACGGTATTTTCGCGTAACCGAGG 1365  
QY 1019 ACG 1021  
Db 1366 ATG 1368

## RESULT 9

US-10-021-323-9015  
; Sequence 9015, Application US/10021323  
; Publication No. US20040123340A1  
; GENERAL INFORMATION:  
; APPLICANT: Deikman, Jill  
; APPLICANT: Feng, Paul C.C.  
; APPLICANT: Fincher, Karen L.  
; APPLICANT: Ziegler, Todd E.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(52274)B  
; CURRENT APPLICATION NUMBER: US/10/021,323  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/255, 619  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 17880  
; SEQ ID NO 9015



Db 260 GGAATAAGAAAGTGTGACCAAGGTTTGGTTCAGGAAGGTTGGCAGGAGTTTCTAGAG 319  
QY 211 CGTTACTCCATTCCGATGTTGTTTCTTTCATTTTATGATATGAAGGAAACTCTGCTTTC 270  
Db 320 CGTTACTATATGTTGTTGCTAGTACTGTTTTCAGATCAGGAAGAAATTCGCTTTC 379  
QY 271 AGCGTCTACATTTTCAATTTATCCACTCTGAGATCAATTTACCATTCCACCGTCTCATG 330  
Db 380 AGTGTTAGTATATTTAATTTGTATACTCGGAATAAACTATCAGACTAATGCCCTCGTT 439  
QY 331 GATTCCGCTCAGACCACTTCAACGCCCGGTTTGTTCAGACCTTGAAGATGAAGA 389  
Db 440 GGTACTCAATACCAATCAGCGAAGAAC---AATATCCGTTTGAACCAACTGAAGATGATGA 495

## RESULT 12

US-10-021-323-7926/c  
; Sequence 7926, Application US/10021323  
; Publication No. US20040123340A1  
; GENERAL INFORMATION:  
; APPLICANT: Deikman, Jill  
; APPLICANT: Feng, Paul C.C.  
; APPLICANT: Fincher, Karen L.  
; APPLICANT: Ziegler, Todd E.  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(52274)B  
; CURRENT APPLICATION NUMBER: US/10/021,323  
; CURRENT FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: US 60/255, 619  
; PRIOR FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 17880  
; SEQ ID NO 7926  
; LENGTH: 602  
; TYPE: DNA  
; ORGANISM: Gossypium hirsutum  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3828-022-Q1-N6-E11  
US-10-021-323-7926

Query Match 9.9%; Score 101.4; DB 17; Length 602;  
Best Local Similarity 73.7%; Pred. No. 4.1e-21;  
Matches 129; Conservative 0; Mismatches 46; Indels 0; Gaps 0;  
QY 847 GAGAAACAATGGCTCTGATGCTCTCAAGCCGGAGAGCCAAATTCAGTCAAGCA 906  
Db 578 GGAGACAATGGCCCAATCGATGCTGTATAGAGTGGCAAGCCAAAGTTTAGTCAAGGA 519  
QY 907 TGGTACGAATTCACCTAGAGAACAACTTAGGAGAAGGAGAGCTCTGTGTGTTGAGCTG 966  
Db 518 TGGTACGAGTTTACGTTGGAGATAATTTGGGTGAAGGAGATGCTGTATCTTCGAGCTG 459  
QY 967 CTCAGAACAGAGATTTGTTTGAAGTGACAGCTTTCGAGTCAACGAGTAAG 1021  
Db 458 CTCAGATCAAGGAATTCGCTTAAAGTTACCGTATTTTCGTTAGGGAAGCG 404

## RESULT 13

US-09-938-842A-2973  
; Sequence 2973, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIPT300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2973  
; LENGTH: 2000  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2973

Query Match 9.8%; Score 101; DB 9; Length 2000;  
Best Local Similarity 88.3%; Pred. No. 1.2e-20;  
Matches 121; Conservative 0; Mismatches 15; Indels 1; Gaps 1;  
QY 96 TGAGCTTTCGGTTCGTTGCTGCACTCAGTACTGATGATGCTGATGTTGGCGTGTAGACT 155  
Db 1784 TGAGCTTTCAGTGTGTTTTCACATACCTGATGCTCATGTTT-GCATGTAGGACT 1842  
QY 156 AAGGAAAGCTGACAAACAAATTTGGTTTCAAGATGTTGGCAAGAGCTTTGTTGACCGTTA 215  
Db 1843 AAGGAAAGCCAAACAAATCTGTTTCAAGACGTTGGCAGGAGTTTGTCAACCGTTT 1902  
QY 216 CTCATTTCGATTGGTT 232  
Db 1903 CTCATTTCGATTGGTT 1919

## RESULT 14

US-09-938-842A-2973  
; Sequence 2973, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIPT300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2973  
; LENGTH: 2000  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2973

Query Match 9.8%; Score 101; DB 11; Length 2000;  
Best Local Similarity 88.3%; Pred. No. 1.2e-20;  
Matches 121; Conservative 0; Mismatches 15; Indels 1; Gaps 1;  
QY 96 TGAGCTTTCGGTTCGTTGCTGCACTCAGTACTGATGCTGATGTTGGCGTGTAGGACT 155  
Db 1784 TGAGCTTTCAGTGTGTTTTCACATACCTGATGCTCATGTTT-GCATGTAGGACT 1842  
QY 156 AAGGAAAGCTGACAAACAAATTTGGTTTCAAGATGTTGGCAAGAGTTTGTGACCGTTA 215  
Db 1843 AAGGAAAGCCAAACAAATCTGTTTCAAGACGTTGGCAGGAGTTTGTCAACCGTTT 1902  
QY 216 CTCATTTCGATTGGTT 232  
Db 1903 CTCATTTCGATTGGTT 1919

```

RESULT 15
US-10-739-930-3743/c
; Sequence 3743, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 3743
; LENGTH: 2207
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2207)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: GLYMA-23APR03-CLUSTER63260_1
US-10-739-930-3743

```

Query Match	8.7%	Score 89.6;	DB 18;	Length 2207;
Best Local Similarity	55.9%;	Pred. No. 5.3e-17;		
Matches	170;	Conservative 0;	Mismatches 134;	Indels 0; Gaps 0;
QY	15	CTTCCATAAGTCGATTTTCTCATCCACTATCCCAAGAAAACGCTCGAGGTCCCAATAA	74	
Db	1758	CTTTTTCCAGATCATTAACCCGCTCACATGTTTCATGAAGAAAGCTATGATCCCAATAA	1699	
QY	75	GTTTGTGAGTAAATTTCAAGGATGAGCTTTCGGTTGCTGTGCTTGCACCTCACAGTACCTGATG	134	
Db	1698	GTTTGTGAAGAAATATGGAATAACCGTTACAAAATACTCTATTTCCTGAAGACTCCAAATGG	1639	
QY	135	TCATGTTTGGGTGTAGGACTAAGGAAAGCTGACACAAAAATTTGGTTTCAAGATCGTTG	194	
Db	1638	TGCTGTAATGGAANAATGATTTTGA AAAACGTGACGCTGAAAATATGGTTTCAAAGGGATG	1579	
QY	195	GCAAGAGTTTGTGACCGTTACTCCATTCGCATTGGTTATCTTTTCATTTTATAGATATGA	254	
Db	1578	GAAAGAGTTTGCAGAGTATCACTCTCTAGCTCATGGCCATCTTTTGGTTTTCAGATGGGA	1519	
QY	255	AGGAAACTCTGCCCTTCAGCGCTCTACATTTTCAATTTATCCACTCTGAGATCAATTAACA	314	
Db	1518	TGTAACCTTCTCAITTTTCAGGTACACATCTTTGATTTGAGTGGCCTTAGAGATTGAGTACCC	1459	
QY	315	TTCC	318	
Db	1458	TACC	1455	

Search completed: December 30, 2004, 17:28:09  
Job time : 664 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 30, 2004, 09:53:25 ; Search time 3904 Seconds  
(without alignments)  
9585.964 Million cell updates/sec

Title: US-10-088-187A-10\_COPY\_269\_1295

Perfect score: 1027  
Sequence: 1 atgcacgcctttcttcca.....agtcacgagtagctgctgaa 1027

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:\*

1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_hic: \*  
4: gb\_est3: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_gsa1: \*  
9: gb\_gsa2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1025.4	99.8	1307	3	CNS0A70B
2	1013.4	98.7	1441	3	BN822585 Arabidops
3	996.6	97.0	1375	3	BN823300 Arabidops
4	470.8	45.8	701	6	BN824523 Arabidops
5	441	42.9	1420	3	CD822418 BN25.045C
6	441	42.9	1490	3	BN816331 Arabidops
7	437.8	42.6	1346	3	BN816536 Arabidops
8	410	39.9	880	3	BN841843 Arabidops
9	401.8	39.1	589	1	AV540912 Arabidops
10	396.2	38.6	780	7	CK120978 204014.p1
11	387	37.7	591	1	AI996485 70166887
12	382.6	37.3	1041	5	BN835682 BN835682
13	380.2	37.0	645	6	CD823066 BN25.047I
14	362.6	35.3	906	7	CD823066 GR_Eb003
15	349.8	34.1	612	6	CD813712 BN15.020L
16	330.2	32.2	785	8	BZ061279 lki10h12
17	329.2	32.1	721	8	BH969696 odd4e04
18	317.6	30.9	834	7	COI08208 GR_Eb003
19	317	30.9	317	1	AV526843 AV526843
20	308.6	30.0	669	1	AV821829 AV821829
21	301.4	29.3	614	1	AV539303 AV539303
22	283.2	27.6	834	7	CO094348 GR_Eal6G
23	273.2	26.6	726	5	BN835627 BN835627
24	267.6	26.1	841	7	CO132469 GR_Eb45F

25	255	24.8	809	7	CO099865
26	255	24.8	872	7	CO128357
27	246.6	24.0	591	6	CA781580 024G12AF
28	236	23.0	453	2	AW705298 ek59c02.Y
29	234.8	22.9	821	7	CO108207 GR_Eb003
30	234.8	22.9	822	7	CO108219 GR_Eb003
31	231.6	22.6	712	7	CO132468 GR_Eb45F
32	227.6	22.2	471	1	AI992821 701493711
33	225.4	21.9	870	4	BM360659 GA_Ea003
34	225	21.9	458	7	T21005 3013.Lambda
35	224.8	21.9	748	7	CO084070 GR_Ea49H
36	224.6	21.9	537	4	BM143791 ga748910.
37	224	21.8	421	1	AV530494 AV530494
38	221.6	21.6	636	6	CD038665 UTPP1010
39	221.2	21.5	582	2	AW187216 BN1GH112
40	219.8	21.4	617	4	BM814806 EST592900
41	219.8	21.4	631	1	AJ504308 AJ504308
42	216.6	21.1	624	2	AW584452 N10302e
43	216.2	21.1	668	5	BQ402703 GA_Ed005
44	216.2	21.1	690	4	BG441541 GA_Ea001
45	216	21.0	414	1	AV533100 AV533100

## ALIGNMENTS

RESULT 1  
CNS0A70B  
LOCUS  
DEFINITION  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSJTFB502E05 of Flowers and buds of strain col-0 of Arabidopsis thaliana (thale cress).  
ACCESSION  
BX822585  
VERSION  
HTC; GSLT\_cDNA  
KEYWORDS  
Arabidopsis thaliana (thale cress)  
SOURCE  
Arabidopsis thaliana  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE  
1 (bases 1 to 1307)  
Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.  
TITLE  
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
JOURNAL  
Unpublished  
REFERENCE  
2 (bases 1 to 1307)  
Genoscope.  
AUTHORS  
Direct Submission  
TITLE  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr - Web : www.genoscope.cns.fr)  
JOURNAL  
The sequences are based on single pass reads.  
COMMENT  
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.  
URV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/banque\_Projet\_EF/Full length  
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.  
Location/Qualifiers  
1. 1307  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"



Qy	61	AGGTC	CCAGATAAGATTG	TGAGTAAAT	CAAGATGAGCTTTCGGTTGCTGTGGCACTC	120
Db	292	AGGTC	CCAGATAAGATTG	TGAGTAAAT	CAAGATGAGCTTTCGGTTGCTGTGGCACTC	351
Qy	121	ACAGT	ACTGATGTCATGTTTGCGGTGTAGGACTAAGGAAGCTGACAA	CAAAATTTGG	180	
Db	352	ACAGT	ACTGATGTCATGTTTGCGGTGTAGGACTAAGGAAGCTGACAA	CAAAATTTGG	411	
Qy	181	TTTCA	GATGGTTGGCAAGAGTTGTTTGA	CGGTTACTCCA	TTGCGATTTGTTTCTTTT	240
Db	412	TTTCA	GATGGTTGGCAAGAGTTGTTTGA	CGGTTACTCCA	TTGCGATTTGTTTCTTTT	471
Qy	241	ATTTTT	TAGATATGAAGAACTCTGCGCTTCAGCGTCTACATTTTCA	ATTTATCCC	ACTCT	300
Db	472	ATTTTT	TAGATATGAAGAACTCTGCGCTTCAGCGTCTACATTTTCA	ATTTATCCC	ACTCT	531
Qy	301	GAGAT	CAATTAACCATTCACCGGTCTCATGGATTCCGCTCACAA	CACTTTCAA	ACGGGCC	360
Db	532	GAGAT	CAATTAACCATTCACCGGTCTCATGGATTCCGCTCACAA	CACTTTCAA	ACGGGCC	591
Qy	361	CGTTT	TGTTGAAGACCTTGAAGATGAAGATGCGAGGTCA	CTTTTCTTTCTTCTGTGTAC	420	
Db	592	CGTTT	TGTTGAAGACCTTGAAGATGAAGATGCGAGGTCA	CTTTTCTTTCTTCTGTGTAC	651	
Qy	421	CCAT	CACCACTTCTCTGAGTCTACAGTACACGACCAACAAAGGGTATGCTAGTTCAGCCATC	480		
Db	652	CCAT	CACCACTTCTCTGAGTCTACAGTACACGACCAACAAAGGGTATGCTAGTTCAGCCATC	711		
Qy	481	CAAA	CTTTGTTTCACTGACACCACTTAAAGCTGAAGAGCCAA	CGCCAAACCCCAAAATACCT	540	
Db	712	CAAA	CTTTGTTTCACTGACACCACTTAAAGCTGAAGAGCCAA	CGCCAAACCCCAAAATACCT	771	
Qy	541	AAAA	GAGAGGGAGGAAGAAAAATGCTGATCCTGAGGAATAA	AACTCATCAGCTCCG	600	
Db	772	AAAA	GAGAGGGAGGAAGAAAAATGCTGATCCTGAGGAATAA	AACTCATCAGCTCCG	830	
Qy	601	CGAG	ATCATCATCCAGAGAACCGTTCAAAGTTCTACGAGAGT	CTTCTGCGAGAAAGAGA	660	
Db	831	CGAG	ATCATCATCCAGAGAACCGTTCAAAGTTCTACGAGAGT	CTTCTGCGAGAAAGAGA	890	
Qy	661	ACCG	TGATGTCAGAAAGAGAGAGAGCCATCAATGCAGCCAA	AAACGGTTCGAACCAACA	720	
Db	891	ACCG	TGATGTCAGAAAGAGAGAGAGCCATCAATGCAGCCAA	AAACGGTTCGAACCAACA	950	
Qy	721	AAC	CTTTCTTTCAGAGTGGTTCTGCGACCATCTTATCTATACAGAGT	TCGATCATGTAT	780	
Db	951	AAC	CTTTCTTTCAGAGTGGTTCTGCGACCATCTTATCTATACAGAGT	TCGATCATGTAT	1010	
Qy	781	CTT	CTTCTCGGTTTGTCTGAGAGTACTTAAGTGGGATCTCCG	GGTTTCATCAAAGTCCAG	840	
Db	1011	CTT	CTTCTCGGTTTGTCTGAGAGTACTTAAGTGGGATCTCCG	GGTTTCATCAAAGTCCAG	1070	
Qy	841	CTT	CGGAGAAACAAATCGCCTGTTTCGATGTCTTACAA	AGCCGGGAGAGCAAAATCAGT	900	
Db	1071	CTT	CGGAGAAACAAATCGCCTGTTTCGATGTCTTACAA	AGCCGGGAGAGCAAAATCAGT	1130	
Qy	901	CAAG	GATGGTACGNAATTCATCTTAGAGAA	CAACTTAGGAGAGTCTGTGTGTTT	960	
Db	1131	CAAG	GATGGTACGNAATTCATCTTAGAGAA	CAACTTAGGAGAGTCTGTGTGTTT	1190	
Qy	961	GAG	CTGCTCAGAACCCAGAGATTTGTTTTCGAAAGTGA	CAGCTTTTCAGTCAACGAGTAC	1020	
Db	1191	GAG	CTGCTCAGAACCCAGAGATTTGTTTTCGAAAGTGA	CAGCTTTTCAGTCAACGAGTAC	1250	
Qy	1021	GTCT	GA	1027		
Db	1251	GTCT	GA	1257		

SECRET

### RESULTS 3

CNS0A6SO

## LOCUS

## DEFINITION

GSLTPGH542B04 of Hormone Treated Callus of strain col-0 of Arabidopsis thaliana (thale cress).  
BX824523  
1 GI:42465662  
HTC; GSLT CDNA  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eutrosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 1375)  
Castelli,V., Aury,J.M., Jaillon,O., Wincher,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
Unpublished  
2 (bases 1 to 1375)  
Genoscope.  
Direct Submission  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr]  
- Web : www.genoscope.cns.fr  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Caetelli V., Aury J.M., Jaillon O., Wincher P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full length  
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.  
Location/Qualifiers  
1. .1375  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLTPGH542B04"  
/tissue\_type="Hormone Treated Callus"  
/plasmid="PCMVSPORT 6"  
complement(1..1375)  
/gene="At3g18990"  
ORIGIN  
Query Match 97.0%; Score 996.6; DB 3; Length 1375;  
Best Local Similarity 98.1%; Pred. No. 1.8e-288;  
Matches 1008; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
QY 1 ATGCCACGCGCCCTTCCTCCATAAGTTGATTTCTCATCCACTATCCAGAAAACGTCG 60  
Db 225 ATGCCACGCGCCCTTCCTCCATGAGTTGATTTCTCATCCACTATCCAGAGGACGTCG 284  
QY 61 AGGGTCCCAAGATAAGTTGTGAGTAAATTCAGAGATGAGCTTCGTTGCTGTTCACATC 120  
Db 285 ATGGTCCCAAGATAAGTTGTGAGTAAATTCAGAGATGAGCTTCGTTGCTGTTCACATC 344  
QY 121 ACAGTACCTGATGGTCATCTTTGGCGGTGTAGGACTTAAGAAAGCTACACAAAAATTTGG 180  
Db 345 ACAGTACCTGATGGTCATCTTTGGCGGTGTAGGACTTAAGAGAGCTGACACAAAAATTTGG 404  
QY 181 TTTCAAGATGGTTGGCAAGATTTGTTGACCGGTACTCCATTCGCATTCGTTATCTTTTG 240  
Db 405 TTTCAGATGGTTGGCAAGATTTGTTGACCGGTACTCCATTCGCATTCGTTATCTTTTG 464  
QY 241 ATTTTTAGATATGAAGAAACTTCGCTTCAGCGTCTACATTTTCAATTTATCCCACTCT 300  
Db 465 ATTTTTAGATATGAAGAAATCTGCTTCAGCGTCTACATTTTCAATTTATCCCACTCT 524

```
QY 301 GAGATCAATTTACATTCACCGGCTCTCATGATTCGGCTCAACACCACTTCAAAACGGCC 360
DB 525 GAGATCAATTTACATTCACCGGCTCTCATGATTCGGCTCAACACCACTTCAAAACGGCC 584
QY 361 CGTTTGGTTGAAGACCTTGAAGATGAAGATGCGAGGTCATCTTTCCTTCTCTGTGTAC 420
DB 585 CGTTTGGTTGAAGACCTTGAAGATGAAGATGCGAGGTCATCTTTCCTTCTCTGTGTAC 644
QY 421 CCATCACCACCTTCCTGAGTCTACACTACCAGCCAAAGGATGCTAGTTCAGCCATC 480
DB 645 CCATCACCACCTTCCTGAGTCTACACTACCAGCCAAAGGATGCTAGTTCAGCCATC 704
QY 481 CAAACCTTGTTCACCTGGACCACTTAAAGCTGAAGAGCCAAAGCCCAACCCCAAAATACCT 540
DB 705 CAAACCTTGTTCACCTGGACCACTTAAAGCTGAAGAGCCAAAGCCCAACCCCAAAATACCT 764
QY 541 AAAAGAGAGGAGGAGGAAGAAATGCTGATCTCTGAGGAAATAAATCTCATGACTCG 600
DB 765 AAAAGAGAGGAGGAGGAAGAAATGCTGATCTCTGAGGAAATAAATCTCATGACTCG 824
QY 601 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGAGTCTTCTGCGAGAAAGAGA 660
DB 825 CGAGATGATGATCCAGAGAACCGTTCAAAGTTCTACGAGAGTCTTCTGCGAGAAAGAGA 884
QY 661 ACCGTGACTGCAGAGAAAGAGAGAGAGCCATCAATGCAGCCCAAAACGTTCCGAACCAACA 720
DB 885 ACCGTGACTGCAGAGAAAGAGAGAGAGCCATCAATGCAGCCCAAAACGTTCCGAACCAACA 944
QY 721 AACCTTCTTTCAGAGTGGTCTCGACATCTATCTATACAGAGGTTGATCATGTAT 780
DB 945 AACCTTCTTTCAGAGTGGTCTCGACATCTATCTATACAGAGGTCGATCATGTAT 1004
QY 781 CTTCTCTTCTGGGTTGCTGAGAGTACCTAAGTGGGATCTCCGGTTCATCAAGTCCAG 840
DB 1005 CTTCTCTTCTGGGTTGCTGAGAGTACCTAAGTGGGATCTCCGGTTCATCAAGTCCAG 1064
QY 841 CTTGCGGAGAAACAAATGGCTCTTCGATGTCTCTACAAAGCGGAGAGCCAAATTCAGT 900
DB 1065 CTTGCGGAGAAACAAATGGCTCTTCGATGTCTCTACAAAGCGGAGAGCCAAATTCAGT 1124
QY 901 CAAGATGATGATCAATTCATCTAGAGAACAACTTAGAGAGAGAGAGTCTGTGTGTT 960
DB 1125 CAAGATGATGATCAATTCATCTAGAGAACAACTTAGAGAGAGAGAGTCTGTGTGTT 1184
QY 961 GAGCTGCTCAGAACAGAGATTTCTGTTTGAAGTGACAGCCTTTCGAGTCAACGAGTAC 1020
DB 1185 GAGCTGCTCAGAACAGAGATTTCTGTTTGAAGTGACAGCCTTTCGAGTCAACGAGTAC 1244
QY 1021 GTCTGAA 1027
DB 1245 GTCTGAA 1251
```

```
RESULT 4
LOCUS CD822418
DEFINITION BN25.045C06F020108 BN25 Brassica napus cDNA clone BN25045C06, mRNA
sequence.
ACCESSION CD822418
VERSION CD822418.1 GI:32504358
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 701)
Genoplante.
AUTHORS a major partnership french program in plant genomics
TITLE Genoplante.
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
```

93, rue Henri Rochefort 91025 EVRY CEDEX France

Tel: 33 1 69 47 54 00

Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french plant genomics programme 'Genoplante' (<http://www.genoplante.com>) and <http://genoplante-info.infobiogen.fr>.

#### FEATURES

source  
1..701  
/organism="Brassica napus"  
/mol\_type="mRNA"  
/cultivar="Jet Neuf"  
/db\_xref="taxon:3708"  
/clone="BN25045C06"  
/tissue\_type="seed"  
/clone\_id="BN25"

#### ORIGIN

Query Match 45.8%; Score 470.8; DB 6; Length 701;  
Best Local Similarity 88.0%; Pred. No. 4.3e-130;  
Matches 537; Conservative 0; Mismatches 67; Indels 6; Gaps 2;  
QY 1 ATGCCAGCCCTTCTTCCATAAGTTGATTTTCTCATCCACTATCCAAAGAAAAGCTGTG 60  
DB 92 ATGCCAGCCCTTCTTCCACAGCTGATTTTCTCATCCACTATCCAAAGAAAAGCTGTG 151  
QY 61 AGGGTCCAGATAAGTTTGTGAGTAAATTTCAAGGATGAGCTTTCGGTTGCTGTTCACCTC 120  
DB 152 AGAGTTCCTGATAGTTTGTGAGTAAATTTCAAGGACGAGCTATCGTTGCGGTTGCTCTC 211  
QY 121 ACAGTACCTGATGTCATGTTTGGCGTGTAGGACTAGGAAAAGCTG---ACAACAATAAT 177  
DB 212 ACAGTACCTGATGTCATGTTTGGCGTGTAGGACTAAGGAAAAGCTGACACACACAAAATC 271  
QY 178 TGGTTTCAAGATGTTGSCAAGAGTTTGTTCACCGTTACTCCATTCGCAATTTGTTATCTT 237  
DB 272 TGGTTTCAAGATGTTGSCAAGAGTTTGTTCACCGTTACTCCATTCGCAATTTGTTATCTT 331  
QY 238 TTGATTTTATGATATGAGGAAACTCTGCGCTTCAGCGCTCTACATTTTCAATTTATCCAC 297  
DB 332 TTGATTTTATGATATGAGGAAACTCTGCGCTTCAGCGCTCTACATTTTCAATTTATCCAC 391  
QY 298 TCTGAGATCAATTAACCATTCACCGCTCTATGATTCGCTCACAACCACTTCAAAAGC 357  
DB 392 TCCGCGATCAACTACCACTCCACCGCTCTATGATTTCTGCACACCAACCACTTCAAAAGC 451  
QY 358 GCCGTTTGTGTTGAAGACCTTGAAGATGAAGATGCGAGGTCATCTTTCCTTCTCTCTG 417  
DB 452 GCGGTTTGTGTTGAAGACCTTGAAGATGAAGATGTTGAGGTTGTTCAACCTTCTCTCTG 511  
QY 418 TACCCATCACCACCTTCTCTGAGTCTACAGTACCAGCCAAACAAAGGATGCTAGTTTCAGCC 477  
DB 512 TACCCATCACCAGCATCTCTGAGACTACTGCGCACGCTAATAAAGGCGATAGTTTCAGCT 571  
QY 478 ATCCAAACCTTGTTCACCTGGACCA---GTTAAAGCTGAAGAGCCACGCCAACCCCAAAA 534  
DB 572 ATCCAGAGCTTCTTCTGCTGAACCACTGTTAAAGCTGAAGAGACAAACCAACCCCAAAA 631  
QY 535 ATACTTAAAGAGAGGAGGAGGAAGAAAATGCTGATCTCTGAGGAATTAACCTATCA 594  
DB 632 GTTCTTAAAGAGAGGAGGAGGAAGAACCGCTGATCTCTGAGGAATTAATTCATCT 691  
QY 595 GCTCCGCGAG 604  
DB 692 GCTCCAGAG 701

#### RESULT 5

##### CNSOAD3Z

##### LOCUS

##### DEFINITION

##### ACCESSION

CNSOAD3Z 1420 bp mRNA linear HTC 06-FEB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLTPGH45ZC05 of Hormone Treated Callus of strain col-0 of  
Arabidopsis thaliana (Chale cress).

VERSION BX16391.1 GI:42473554  
KEYWORDS HTC; GSLT cDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 1420)  
AUTHORS Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,  
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,  
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.  
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1420)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
COMMENT The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,  
URV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
FEATURES  
source  
1..1420  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLTPGH452C05"  
/tissue\_type="Hormone Treated Callus"  
/plasmid="pCMVSPORT 6"  
complement(1..1420)  
/gene="At1g49480"  
Query Match 42.9%; Score 441; DB 3; Length 1420;  
Best Local Similarity 73.5%; Pred. No. 5.4e-121;  
Matches 646; Conservative 0; Mismatches 195; Indels 38; Gaps 5;  
QY 151 GGACTAAGGAAGCTGACACAAATTTGGTTTCAAGATGGTTGGCAAGAGTTCTTCTGAC 210  
DB 336 GGACTAAGGAAGCTGACACAAATTTGGTTTCAAGATGGTTGGCAAGAGTTCTTCTGAC 395  
QY 211 CGTTACTCCATTCGCGATTGGTTATCTTTGATTTTATAGATGAAGGAACCTCGCTTC 270  
DB 396 CGTTTCTCCATTCGCGATTGGTTATCTTTGATTTTATAGATGAAGGAACCTCGCTTC 451  
QY 271 AGCGTCTCAATTTCAATTTATCCCACTCTGAGATCAATTTACCATTCACCGGTTCTGATG 330  
DB 452 ATCTCTCCACACTCTGAGATCAACACC-----ATTCTAGTAGTGAAGCTCTTTATGCAATG 507  
QY 331 GATTCGGCTCAACACCTTT---GAAACGGCCCGTTTGTGTGAAGACCTTGAAGATGAA 387  
DB 508 GATTCGGCACAGAAATCAGTTCAACAAACGTCGCTGATTTGTTTGAAGATCTCTGAACTCAA 567  
QY 388 GATGCGGAGGTCTATCTTCTCTTCTCTGTTGATCCCATCACCATTCTCTGAGTCTACGTA 447  
DB 568 GATGCTAAGGTCTATCTTCTCTTCTCTGTTGATCCCATCACCATTCTCTGAGTCTACGTA 609  
QY 448 CCAGCCAAACAAGGGTATGCTAGTTGAGCTCATCCAAACCTGTTTCTGAGTCTGAGTCTGAA 507

DB 610 CCAAGTGAATAAAGGTTATGGCGGTTCTATAGCCATCAAAAGCTTTTCAAGAATCTTAAA 669  
QY 508 GCTGAAGAGCCAAACGCCCAACCCAAAAATACCTTAAAGAGAGAGGAGGAGGAAGAAAT 567  
DB 670 GCTGAAGA-----AACTCCCAAGTACTTAAAGAGAGAGGAGGAGGAAGAAAT 720  
QY 568 GCTGATCTCTAGGAAATAAATCAATCATCAGCTCCGCGAGATGATATCCAGAGAACCGTTCA 627  
DB 721 CCTAATCCGAGGAAGTAAATCTTTCAACTCCCGTGGAGATGATCTCAGAGAACCGCTCA 780  
QY 628 AAGTTCTACGAGAGTCTTCTCGGAGGAAGAGAACCGTACTGCGAGAGAAAGAGAGAGA 687  
DB 781 AAGTTCTACGAGAGTCTTCTGCTAGAAAGAGAACTGTAACCTGCGAGAGAAAGAGAGAGA 840  
QY 688 GCCATCAATCAGCCCAAAACGTTCCGAAACCAACCAACCTTTCTTCAGAGTGGTTCTGCCA 747  
DB 841 GCCGTCATTCAGCCCAAAACATTCGAAACCAACCAACCTTTCTTCAGAGTGGTTCTGCCA 900  
QY 748 CCATCTATCTATACAGAGTTGCAATCATGATATCTTCTCTGCGTGGTTGCTGAGAGTAC 807  
DB 901 CCATCATATCTATACAGAGTTGCAATCATGATATCTTCTCTGCGTGGTTGCTGAGAGTAC 960  
QY 808 CTAAGTGGATCTCCGGTTTCATCAAACTCCAGCTTCGGGAGAAACAATGCTCTTCA 867  
DB 961 CTAAGTGGATCTTGGTTTCATCAAGCTCCAGCTTCGGTGAAGAAACAATGCGCAGTGG 1020  
QY 868 TGTCTCTACAAAGCGGAGAGCCAAATTCAGTCAAGGATGATGATGATGATGATGATGATGAT 927  
DB 1021 TGCCTCTACAAAGCGGAGAGCTAAGTTTAGCCAGGATGATGATGATGATGATGATGATGAT 1080  
QY 928 AACAACTTAGGAGAGAGAGAGCTGCTGCTGTTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCT 987  
DB 1081 AACAACTTAGGAGAGAGAGAGTGTGTGTGTTGAGTACTCAGAACTCGGGATTCGTT 1140  
QY 988 TTGAAGTGAAGAGCTTTCAGTCAACAGTACGTTCTCA 1026  
DB 1141 CTCGAAGTCAACCGCTTTTCGTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1179  
RESULT 6  
CNSOAD86 1490 bp mRNA linear HTC 06-FEB-2004  
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
DEFINITION GSP132B06 of Hormone Treated Callus of strain col-0 of  
Arabidopsis thaliana (thale cress).  
ACCESSION BX15859  
VERSION BX15859.1 GI:42473387  
KEYWORDS HTC; GSLT cDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 1490)  
AUTHORS Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,  
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,  
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.  
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1490)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
COMMENT The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,  
Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences). 5 prime and 3 prime are assembled with Phrap.  
[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/Full](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full)  
length  
<http://www.genoscope.cns.fr/cgi-bin/ggb?ggb?source=Arabidopsis>.

## FEATURES

Location/Qualifiers  
1..1490  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLTPGH132B06"  
/tissue\_type="Hormone Treated Callus"  
/plasmid="pCMVSPORT 6"  
complement(1..1490)  
/gene="At1g49480"

## ORIGIN

Query Match 42.9%; Score 441; DB 3; Length 1490;  
Best Local Similarity 73.5%; Pred. No. 5.5e-121;  
Matches 646; Conservative 0; Mismatches 195; Indels 38; Gaps 5;  
QY 151 GGACTAAGGAAGCTGACACAAATTTGGTTTCAAGATGGTTGCAAGAGTTTCTTGAC 210  
DB 225 GACTAAGGAAGCCACACAAATCTGGTTTCAAGACGGTTGCGAGAGTTTGTCAAC 284  
QY 211 CGTTACTCCATTCGATGGTTTATCTTTTGTATGATATGAAGAACTCTCCCTTC 270  
DB 285 CGTTTCTCCATTCGGATTGGT---TTCCAGATACAAAGTACAGTCTACATTTCAAATTT 340  
QY 271 AGCGTCTACATTTCAATTTATCCACTCTGAGATCAATTACCATTCCACCGTCTCATG 330  
DB 341 ATCTCCACACTCTGAGATCAACACC---ATTCTAGTAGGAAGCTCTTATGCAATG 396  
QY 331 GATTCGGCTCAACACACTT---CAACGCGCGCTTTGTTTGAAGACCTTGAAGATGAA 387  
DB 397 GATTCGCACAGATCAGTTCAACAAAGTGTCTGATTTTGAAGATCTTGAATCTCAA 456  
QY 388 GATCGCGAGGTCAATCTTCTCTTCTGTGTACCATCACCACTTCTTGAGTCTACGTA 447  
DB 457 GATCGGAAGTCAATTC---CATCGAACCTTGAATCTACTGAA 498  
QY 448 CCAGCCACAAAGGTATGCTAGTTCAGCCATCAAACTGTTTCACTGGACCACTTAA 507  
DB 499 CCAATGAATAAAGTTATGGCGGTTCTACAGCCATCAAAAGCTTTTCAAGAATCTAA 558  
QY 508 GCTGAAGAGCCAAAGCCAAACCAAAATACCTAAAGAGAGGAGGAGGAAGAAAT 567  
DB 559 GCTGAAGA-----AAGCCCAAGGTACTTTAAGAGAGAGGAGGAGGAGGAGGAG 609  
QY 568 GCTGATCTCTAGGAATAAATCATCAGTCCGCGAGATGATGATCCAGAGAACCGTTCA 627  
DB 610 CCTAATCCGCGAAGTAATCTTCAACTCCCGGTGGAGATGATCTCAGAGAACCGCTCA 669  
QY 628 AGTTCTACGAGAGTCTTCTCGAGAGAGAGAACCGTGACTGCGAGAGAGAGAGAGA 687  
DB 670 AAGTTCTACGAGAGTCTTCTGCTAGAAGAGAGACTGTAACCTGACGAGAGGAGAGAGA 729  
QY 688 GCATCAATGACGCAAAAGCTTCAACCAAAACCCCTTTCTTCAAGAGTGGTTCTCGGA 747  
DB 730 GCGTCAATGACGCAAAACATTCGAAACCAATCTTACTTTAGAGTGGTTCTCGGA 789  
QY 748 CCATCTATCTATACAGAGGTTGATCATGATCTTCTTCTTCTGGTTTCTGAGAGTAC 807  
DB 790 CCATCATATCTATACAGAGGTTGATCATGATCTTCTTCTTCTGGTTTCTGAGAGTAC 849  
QY 808 CTAAGTGGATCTCCGGTTTCATCAAGTCCAGCTTGGGAGAAACATGGCTCTTGA 867  
DB 850 CTAAGTGGATATCTGGTTTTCATCAAGCTCCAGCTCGGTGAGAAACATGGCCAGTGG 909  
QY 868 TGTCTCTACAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGATGATTCATCTAGAG 927

DB 910 TGCCTCTACAAAGCAGGAGAGCTAAGTTTAGCAAGGATGATGAGTTCCACTCGAG 969  
QY 928 AACAACTTAGAGAGAGAGAGCTCTGTGTGTTTGTAGCTGCTCAGAACAGAGATTCGTT 987  
DB 970 AACAACTTAGAGAGAGAGATGATGTGTTTGTAGCTACTCAGAACTCGGATTCGTT 1029  
QY 988 TTGAAGTGACAGCTTTTCGAGTCAACGAGTACGTCTGA 1026  
DB 1030 CTCGAAGTCACCGCTTTCGTGTCATGAGTATGTGTA 1068  
CNS0AD6P 1346 bp mRNA linear HTC 06-FEB-2004  
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from Clone  
DEFINITION GSLTPGH542B06 of Hormone Treated Callus of strain col-0 of  
Arabidopsis thaliana (thale cress).  
ACCESSION BX816536  
VERSION BX816536.1 GI:42473594  
KEYWORDS HTC; GSLT cDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
REFERENCE 1 (bases 1 to 1346)  
AUTHORS Castell, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,  
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,  
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.  
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1346)  
AUTHORS Genoscope.  
JOURNAL Direct Submission  
TITLE Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
COMMENT The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,  
Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences). 5 prime and 3 prime are assembled with Phrap.  
[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/Full](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full)  
length  
<http://www.genoscope.cns.fr/cgi-bin/ggb?ggb?source=Arabidopsis>.  
Location/Qualifiers  
1..1346  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLTPGH542B06"  
/tissue\_type="Hormone Treated Callus"  
/plasmid="pCMVSPORT 6"  
complement(1..1346)  
/gene="At1g49480"

FEATURES  
source  
Query Match 42.6%; Score 437.8; DB 3; Length 1346;  
Best Local Similarity 73.3%; Pred. No. 4.8e-120;  
Matches 644; Conservative 0; Mismatches 197; Indels 38; Gaps 5;  
QY 151 GGACTAAGGAAGCTGACACAAATTTGGTTTCAAGATGGTTGCAAGAGTTTCTTGAC 210  
DB 229 GGACTAAGGAAGCCACACAAATCTGTTTCAAGACGGTTGCGAGAGTTTGTCAAC 288

```
QY 211 CGTTACTCCATTCGCAATGGTTATCTTTGATTTTATGATATGAAGAAACTTCGCTTC 270
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 289 CGTTTCTCCATTCGGATTCGT-----TTCAGATACAAAGTTACAGTCTACATTTCAATTT 344
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 271 AGCGTCTACATTTTCAATTTATCCCACTCTGAGATCAATACCATTCACCGGTCTCAG 330
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 345 ATCTCTCACACTGAGATCAACACC-----ATTCTAGTAGTGAAGCTCTTATGCAATG 400
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 331 GATTCGCTCACAAACACTT---CAAACGCGCCGTTTGTGTAAGACCTTGAAGATGAA 387
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 401 GATTCGCGACAGATCAGTTCAACAAAGTCTGATTTGTAAGATCTGTAACCTCAA 460
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 388 GATCGCAGGTCACTTTCTTCTTCTGTGTACCCATCACTTCTCTGAGTCTACAGTA 447
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 461 GATGCTAAGTCAATTC-----CATCGAACCTGAATCTACTGAA 502
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 448 CCAGCCAAACAAGGTATGCTAGTTCAGCATCAACACTTGTCTACTGGACAGTTAA 507
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 503 CCAGTGAATAAAGTTATGCGGTTCTACAGCCATCAAGCCTTTTCAAAGATCTAAA 562
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 508 GCTGAAGAGCCACGCCAACCCCAAAATACCTTAAAGAGAGGAGGAGGAAGAAAT 567
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 563 GCTGAAGA-----AAGGCCAAGGTACTTAAAGAGAGAGGAGGAAGAAAGAT 613
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 568 GCTGATCTCTGAGGAATAAACTCATCAGCTCCGCGAGATGATGCCAGAAACCGTTCA 627
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 614 CCTAATCCCGAGGAGTAACTCTCACTCCCGTGGAGATGACTCAGAGAACCGCTCA 673
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 628 AGTTCTACAGAGTCTCTCGGAGAAAGAGAACCGTGAATGCGAGAGAAAGAGAGAGA 687
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 674 AAGTTCTACGAGAGTTCTTCTGTAGAAAGAGAACTGTAACTGCAGAGGAAGAGAGAGA 733
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 688 GCATCAATGCAGCCAAAAGTTTGAACCAACAAACCTTTCTTCAAGTGTGTTCTGCA 747
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 734 GCGTCAATGCAGCCAAAACATTCGAACCAACAAATCCTTACTTATAGTGTGTTCTGCA 793
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 748 CCATCTATCTATACAGAGGTGTCATCATGTATCTTCTTCTGCGTTTCTGAGAGTAC 807
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 794 CCATCATATCTATACAGAGGTGTCATCATGTACTTCCATCTGGTTTCTGAGAAATAC 853
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 808 CTAAGTGGATCTCCGGTTTCATCAAAGTCCAGTTCGCGAGAAACAAATGGCTGTGCA 867
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 854 CTAAGTGGATATCTGTTTTCATCAAGCTCCAGTCCGTTGAGAAACAAATGGCCAGTGAG 913
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 868 TGTCTCTACAAAGCCGGGAGACCAATTCAGTCAAGGATGCTACGAATTCATCTCAGAG 927
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 914 TGCTCTCAAAAGCGGGAGAGAGCTAAGTTTAGCCAAAGGATGGTATGAGTTCACTCGAG 973
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 928 AACAACTTAGGAGAGAGAGAGCTGTGTGTGTTGAGTCTCTCAGAACCAAGATTCGTT 987
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 974 AACATATAGCGAGAGAGATGATGTGTGTTGAGTACTCAGAACTCGGATTCGTT 1033
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 988 TTGAAGTACAGCCCTTCGAGTCAACAGTAGTCTGA 1026
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1034 CTCGAAGTCAACCGCTTCGTGTCAATGATGATGTGA 1072
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 8

```
CNS09Y69 880 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSL7S76ZC03 of adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
EX841843
VERSION BX841843.1 GI:42454500
KEYWORDS HTC; GSLT_cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 880)
```

## AUTHORS

Castelli, V., Aury, J.M., Jallou, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, P., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

## JOURNAL

Unpublished 2 (bases 1 to 880)

## REFERENCE

Genoscope.

## AUTHORS

Direct Submission Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr) - Web : www.genoscope.cns.fr

## JOURNAL

The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.

## COMMENT

URG V INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.

## length

http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis. Location/Qualifiers 1..880 /organism="Arabidopsis thaliana" /mol\_type="mRNA" /strain="Col-0" /db\_xref="taxon:3702" /clone="GSL7S76ZC03" /tissue\_type="Adult vegetative tissue" /plasmid="pCMVSPORT\_6" complement(1..880) /gene="Atlg49480"

## gene

## ORIGIN

Query Match 39.9%; Score 410; DB 3; Length 880; Best Local Similarity 72.7%; Pred. No. 9.9e-112; Matches 640; Conservative 0; Mismatches 200; Indels 40; Gaps 7;

## QY

151 GGACTAAGGAAGCTGACAAACAAATTTGGTTTCAAGATGGTTGGCAAGAGTTGTTGAC 210

## Db

1 GGACTAAGGAAGCCAAACAAATCTGTTTCAAGACGGTTGGCAGAGTTGTTCAAC 60

## QY

211 CGTTACTCCATTCGCAATGGTTATCTTTGATTTTATGATATGAAGAAACTTCGCTTC 270

## Db

61 CGTTTCTCCATTCGGATTCGT---TTCAGATACAAAGTTACAGTCTACATTTCAATTT 116

## QY

271 AGCGTCTACATTTTCAATTTATCCCACTCTGAGATCAATACCATTCACCGGTCTCAG 330

## Db

117 ATCTCTCCACACTCTGAGATCAACCACTTCTAGTAGTGAAGCTCTTATGCG-----AA 172

## QY

331 GATTCGCTCACAAACACTT---CAAACGCGCCGTTTGTGTAAGACCTTGAAGATGAA 387

## Db

173 GATTCGCGACAGATCAGTTCAACAAAGTGTCTCATTTGTTGAGATCTCTGAACCTCAA 232

## QY

388 GATCGCAGGTCACTTTCTTCTTCTGTGTATCCCATCACTTCTCTGAGTCTACAGTA 447

## Db

233 GATGCTAAGTCAATTCATTC-----CATCGAACCTGAATCTACTGAA 274

## QY

448 CCAGCCAAACAAGGTATGCTAGTTCAGGCATCCAAACCTTGTCTCACTGGACAGTTAA 507

## Db

275 CCAGTGAATGAGGGTTATGGCGGTTTACAGCCATCCAAAGCTTTTTCAAAGAACTCAA 334

## QY

508 GCTGAAGAGCCAAACGCCAACCCCAAAATACCTTAAAGAGAGGAGGAGGAAGAAAT 567

## Db

335 GCTGAAGAG-----ACGCCAAGGTACTTAAAGAGAGGAGGAAGGAAGAAT 385

## QY

568 GCTGATCTCTGAGGAATAAACTCATCAGCTCCGCGAGATGATGATCCAGAGAACCGTTCA 627

Db 386 CTAATCCGAGAGAAAGTAAACTTTCACTCCCGTGGAGATGACTCAGAGAACCGCTCA 445  
QY 628 AAGTTCTACGAGAGTCTTCTCGGAGAAAGAGAACCGTACTGC-AGAAGAAAGAGAG 686  
Db 446 AAGTTCTACGAGAGTCTTCTCGGAGAAAGAGAACCGTACTGC-AGAAGAAAGAGAG 505  
QY 687 ACCCATCAATGACGAGCAAAACGTTTCGAAACCAACAAACCTTTCTTTCAGAGTGTCTCGG 746  
Db 506 AGCGGTCAATGACG-CAAAACATTGCAACCAACAAATCCTTACTTTCAGAGTGTCTCGG 564  
QY 747 ACCATCTATCTATACAGAGTGTGATCATGATATCTTCTTCGAGTGTCTCGGTTGCTGAGA 806  
Db 565 ACCATCATATCTATACAGAGTGTGATCATGATATCTTTCGAGTGTCTCGGTTGCTGAGA 624  
QY 807 CTTAAGTGGGATCTCCGGTTCATCAAGTCCAGTTCGGGAGAAACAAATGCGCTGTTCG 866  
Db 625 CTTAAGTGGGATCTCCGGTTCATCAAGTCCAGTTCGGGAGAAACAAATGCGGAGTGTGAG 684  
QY 867 ATGTCTCTACAAAGCGGGAGAGCAAAATTCAGTCAAGGATGTGATGATGATGATGATGAT 926  
Db 685 GTGCTCTACAAAGCGGGAGAGTAAAGTTTACGAGGATGATGATGATGATGATGATGATGAT 744  
QY 927 GAACAACTTAGAGAGAGGAGAGT 986  
Db 745 GAACGATATAGCGAGGAGAGT 804  
QY 987 TTTGAAAGTGACAGCTTTTCGAGTCAACGAGTACGCTGTA 1026  
Db 805 TCTCGAAGTACCGGCTTTTCGTTCAATGATGATGTGTGA 844

RESULT 9  
LOCUS AV540912/c 589 bp mRNA linear EST 20-FEB-2004  
DEFINITION AV540912 Arabidopsis thaliana roots Columbia Arabidopsis thaliana  
CDNA clone RZ157509F 3', mRNA sequence.  
ACCESSION AV540912  
VERSION AV540912.1 GI:8702670  
KEYWORDS EST.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE 1 (bases 1 to 589)  
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.  
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries  
JOURNAL DNA Res. 7 (3), 175-180 (2000)  
MEDLINE 20363093  
PubMed 10907847  
COMMENT Contact: Erika Asamizu  
The First Laboratory for Plant Gene Research  
Kazusa DNA Research Institute  
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES  
source location/Qualifiers  
1..589  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="RZ157509F"  
/tissue\_type="roots"  
/clone\_lib="Arabidopsis thaliana roots Columbia"  
/note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2: XhoI"

## ORIGIN

Query Match 39.1%; Score 401.8; DB 1; Length 589;  
Best Local Similarity 99.5%; Pred. No. 2.5e-109;  
Matches 403; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 623 GTTCAAAGTTCTACGAGAGTCTTCTCGGAGAAAGAGAACCGTACTGCAGAGAAAGAG 682  
Db 589 GTTCAAAGTTCTACGAGAGTCTTCTCGGAGAAAGAGAACCGTACTGCAGAGAAAGAG 530  
QY 683 AGAGAGCCATCAATGACGAGCAAAACGTTTCGAAACCAACAAACCTTTCTTTCAGAGTGTGTT 742  
Db 529 AGAGAGCCATCAATGACGAGCAAAACGTTTCGAAACCAACAAACCTTTCTTTCAGAGTGTGTT 470  
QY 743 TCGGACCATCTATCTATACAGAGTGTGATCATGATATCTTCTTCGAGTGTGCTGAGA 802  
Db 469 TCGGACCATCTATCTATACAGAGTGTGATCATGATATCTTCTTCGAGTGTGCTGAGA 410  
QY 803 AGTACTAAGTGGGATCTCCGGTTCATCAAGTCCAGTTCGGGAGAAACAAATGCGCTG 862  
Db 409 AGTACTAAGTGGGATCTCCGGTTCATCAAGTCCAGTTCGGGAGAAACAAATGCGCTG 350  
QY 863 TTGATGTCTCTACAAAGCGGGAGAGCAAAATTCAGTCAAGGATGTGATGATGATGATGATGAT 922  
Db 349 TTGATGTCTCTACAAAGCGGGAGAGCAAAATTCAGTCAAGGATGTGATGATGATGATGATGAT 290  
QY 923 TAGAGAACAACTTAGGAGAGGAGAGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 982  
Db 289 TAGAGAACAACTTAGGAGAGGAGAGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 230  
QY 983 TCGTTTTGAAAGTGACAGCTTTTCGAGTCAACGAGTACGCTCTGAA 1027  
Db 229 TCGTTTTGAAAGTGACAGCTTTTCGAGTCAACGAGTACGCTCTGAA 185

RESULT 10  
LOCUS CK120978 780 bp mRNA linear EST 01-JUN-2004  
DEFINITION 204014.p1 AtM1 Arabidopsis thaliana cDNA clone MPWGP2011014204  
5-PRIME, mRNA sequence.  
ACCESSION CK120978  
VERSION CK120978.1 GI:47831294  
KEYWORDS EST.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE 1 (bases 1 to 780)  
AUTHORS Feilner,T., Immink,R.G.H., Cahill,D.J. and Kersten,B.  
TITLE Generation of a cDNA expression library from Arabidopsis inflorescence meristem  
JOURNAL Unpublished (2003)  
COMMENT Contact: Birgit Kersten  
Plant Protein Chip Group, Department Leirach  
Max-Planck-Institute for Molecular Genetics  
Innestr. 73, D-14195 Berlin, Germany  
Tel: +49(0)30/84131648  
Fax: +49(0)30/84131128  
Email: Kersten@molgen.mpg.de  
Insert Length: 780 Std Error: 0.00  
Plate: 204 row: 0 column: 14  
Seq primer: PQ85 location/Qualifiers  
1..780  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Columbia"  
/db\_xref="GABI:952488"  
/db\_xref="taxon:3702"  
/clone="MPWGP2011014204"  
/tissue\_type="inflorescence meristem"  
/dev\_stage="about one week after bolting"  
/lab\_host="E. coli SCS-1/pSE111"  
/clone\_lib="AtM1"  
/note="Vector: PQ8-3ONAST-attB (AY386205); Site 1: Salt; Site 2: NotI; About 1 week after bolting, cDNA synthesis using SuperscriptTM-system (Invitrogen) with an

FEATURES  
source



Db	352	GGCCTGTTGATGCTCTTACAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGGTACGAT	293
Qy	917	TCATCTCTAGAGAACAACTTAGGAGAGGAGAGCTGTGTGTGTTGAGCTGCTCAGAACCA	976
Db	292	TCATCTCTAGAGAACAACTTAGGAGAGGAGAGCTGTGTGTGTTGAGCTGCTCAGAACCA	233
Qy	977	GAGATTCGTTTGAAGAGGAGAGCTTTCGAGTCAAGAGTACGTCGAA	1027
Db	232	GAGATTCGTTTGAAGAGGAGAGCTTTCGAGTCAAGAGTACGTCGAA	182
RESULT 12			
BX835682/c			
LOCUS			
DEFINITION			
Arabis thaliana Adult vegetative tissue Col-0			
Arabis thaliana cdna clone GSLTSL48ZH11 3PRIM, mRNA sequence.			
ACCESSION			
BX835682			
VERSION			
BX835682.1			
SOURCE			
EST.			
ORGANISM			
Arabis thaliana (thale cress)			
REFERENCE			
Castelli, V., Aury, J.M., Jallou, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.			
TITLE			
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation			
JOURNAL			
Unpublished (2004)			
COMMENT			
Genoscope - Centre National de Sequencage			
BP 191 91006 EVRY cedex - France			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
The sequences are based on single pass reads.			
Life Technologies (a division of Invitrogen) members carried out			
full-length libraries construction : Temple G.			
Genoscope members carried out sequencing and annotation : Castelli			
V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.			
URGV INRA : Clepet C., Caboche M.			
Annotation is based on the June 2003 version of the Arabidopsis			
genome released by MIPS (Munich Information center for Protein			
Sequences).			
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST			
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.			
FEATURES			
source			
1..1041			
/organism="Arabidopsis thaliana"			
/mol_type="mRNA"			
/ecotype="Col-0"			
/db_xref="taxon:3702"			
/clone="GSLTSL48ZH11"			
/tissue_type="Adult vegetative tissue"			
/clone_lib="Arabidopsis thaliana Adult vegetative tissue			
Col-0"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
470; Conservative			
0; Mismatches			
114; Indels			
9; Gaps			
1;			
Qy	434	CTGAGTCTACAGTACCGACCAACAGGGTATGCTAGTTCAGGATCCAAACCTTTTCA	493
Db	876	CTGAAACTACTGACCAAGTGAATAGGTTATGCGGTTCTACAGCCATCCAAAGCTTT	817
Qy	494	CTGAGCAGATTAAAGCTGAAGACCAACGCCCAACCCCAAAATACCTAAAGAGAGGGA	553
Db	816	TCAAAGAAATCTAAAGCTGAAGA-----AAGCCCAAGGTTACTTAAGAGAGGAA	766
Qy	554	GGAAGAGAAAAATGCTGATCTCTGAGGAAATAAATCATCATGCTCCGCGAGATGATC	613
Db	765	GGAGAGAGAGATCTTAATCCCGGAGAGTAAACTCTTCAACACCCGGTGGAGATGACT	706
Qy	614	CAGAGAACCTTCAAGATTTCTACGAGTGTCTTCTGGGAGAAAGAGACCGTACTGTCAG	673
Db	705	CAGAGAACCTTCAAGATTTCTACGAGTGTCTTCTGGTAGAAAGAGAACTGTAACTGAG	646
Qy	674	AGAGAGAGAGAGAGCCATCAATGCGAGCCAAAGGTTCCGAAACCAACAAACCTTTCTTCA	733
Db	645	AGGAGAGAGAGAGAGCCGCTCAATGCGAGCCAAACCAATTCGAAACCAACAAATCTTCTTA	586
Qy	734	GAGTGGTCTTCCGACCATCTCTATATACAGAGGTTGCATCATGTATCTTCTTCTGGGT	793
Db	585	GAGTGGTCTTCCGACCATCTCTATATACAGAGGTTGCATCATGTATCTTCTTCTGGGT	526
Qy	794	TTGCTGAGAGTACCTAAGTGGGATCTCCGGTTCATCAAGTCCAGCTTCCGAGAGAAC	853
Db	525	TTGCTGAGAAATACCTAAGTGGGATCTCTGTTTTCATCAAGTCCAGCTCCGAGAGAAC	466
Qy	854	AATGGCTGTTCCGATGTTCTTACAAAGCCGGAGAGCCAAATTCAGTCAAGGATGGTACG	913
Db	465	AATGGCCAGTGGTGGTCTCTACAAAGCAGGAGAGCTAAGTTTACCAAGGATGGTATG	406
Qy	914	AATTCACCTCTAGAGAAACAACTTAGGAGAGAGAGCTGTGTGTGTTTGGCTCTCAGAA	973
Db	405	AGTTTCACTCTAGAGAAACAACTTAGGAGAGAGAGATGTATGTGTGTTTGGCTCTCAGAA	346
Qy	974	CCAGAGATTTGTTTGAAGTGGACGCTTTCAGTCAAGGATGAGTCTCTGA	1026
Db	345	CTCGGATTTTCTCGAAGTCAACCCCTTTCGTGTCATGATGATGTGTGA	293
RESULT 13			
CD823066			
LOCUS			
DEFINITION			
BN25.047120F020109 BN25 Brassica napus cdna clone BN25047120, mRNA			
ACCESSION			
CD823066			
VERSION			
CD823066.1			
KEYWORDS			
EST.			
SOURCE			
ORGANISM			
Brassica napus (rape)			
REFERENCE			
1 (bases 1 to 645)			
AUTHORS			
Genoplante.			
TITLE			
Unpublished (2003)			
JOURNAL			
COMMENT			
Contact: Genoplante			
Genoplante			
93, rue Henri Rochefort 91025 EVRY CEDEX France			
Tel: 33 1 69 47 54 00			
Fax: 33 1 69 47 54 10			
This sequence has been generated in the framework of the french			
plant genomics programme 'Genoplante' (http://www.genoplante.com			
and http://genoplante-info.infobiogen.fr).			
FEATURES			
source			
1..645			
/organism="Brassica napus"			
/mol_type="mRNA"			
/cultivar="Jet Neuf"			
/db_xref="taxon:3708"			
/clone="BN25047120"			
/tissue_type="seed"			
/clone_lib="BN25"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
412; Conservative			
0; Mismatches			
53; Indels			
0; Gaps			
0;			
Qy	551	GGAGGAGAGAGAAATGCTGATCTCGAGGAAATAAATCATCATGCTCCGCGAGATGATG	610

```

1  GGAGGAGAGAGAGAGAGCGTGTTCCTGAGAGAGTAAACTCATCTGCTCCGAGGGATGATG 60
611 ATCCAGAGAACCCGTTCAAAAGTTCTACGAGAGTGTCTCTGCGAGAAAGAGAACCGTGACTG 670
61  ACCCGAGAGCCGTTCAAAAGTTCTACGAGAGTGTCTCTGCGAGAAAGAGAACCGTTACTG 120
671 CAGAAGAAAGAGAGAGAGCCATCAATGCGAGCCAAAAGCGTTGAAACCAACAAACCCCTTCT 730
121 CAGAGGAAAGAGAGAGAGGGCCATTAATGCGAGCCAAAAGCGTTGAGCCCAACAAACCCCTTCT 180
731 TCAGAGTGTCTCTGCGAGCCATCTATCTATACAGAGGTTGCGATCATGTATCTCTCTCTG 790
181 TCAGAGTGTCTCTGCGAGCCATCTATCTATACAGAGGTTGCGATCATGTATCTCTCTCTG 240
791 GGTTCCTGAGAGTACCTAAAGTGGGATCTCCGGGTTCATCAAGTCCAGCTTGGCGAGA 850
241 GTTTCCTGAGAGTACTTAAGTGGGATCTCCGGGTTCATCAAGTCCAGCTTGGCGAGA 900
851 AACAAATGGCTGTTCGATGTCTCTACAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGT 910
301 AACAGTGGCCGTTGAGATGCTTTTACAAAGCAGGAGAGCCAAATTCAGTCAAGGATGT 360
911 ACGAATTCCTCTAGAGAACCACTTAGGAGAGGAGAGAGTCTGTGTGTTTGAAGTCTCTCA 970
361 ACAGATTCCCTCGGAGAACCACTTAGGAGAGGAGAGTCTGTGTGTTTGAAGTCTCTCA 420
971 GAACAGAGATTTGTTTGAAGTGCAGAGCTTTTCGAGTCAACG 1015
421 GAACAGAGATTTGTTTGAAGTGCAGAGCTTTTCGAGTCAACG 465

```

## RESULT 14

LOCUS CO108220 GR\_Eb0039N17.r GR\_Eb Gossypium raimondii cDNA clone GR\_Eb0039N17 3', mRNA sequence.

## ACCESSION

CO108220

## VERSION

CO108220.1 GI:48806906

## KEYWORDS

EST.

## SOURCE

Gossypium raimondii

## ORGANISM

Gossypium raimondii

## REFERENCE

AUTHORS

Kim.H., Yu.Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C., Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and Wing,R.A.

Title

JOURNAL

COMMENT

Global assembly of Cotton ESTs

Unpublished (2004)

Contact: Rod A. Wing

Arizona Genomics Institute

The University of Arizona

Forbes Building Room 303, Tucson, AZ, 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: http://genome.arizona.edu

Plate: 0039 row: N column: 17.

Location/Qualifiers

1. .906

/organism="Gossypium raimondii"

/mol\_type="mRNA"

/db\_xref="taxon:29730"

/clone="GR\_Eb0039N17"

/tissue\_type="floral"

/dev\_stage="3 to +3 DPA"

/lab\_host="DH10B"

/clone\_lib="GR\_Eb"

/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2: EcoRV; library made by Invitrogen with RNA supplied by Wendle lab. Directional clones into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."

## FEATURES

source

## ORIGIN

Query Match 35.3%; Score 362.6; DB 7; Length 906;  
Best Local Similarity 64.8%; Pred. No. 1.9e-97;  
Matches 587; Conservative 0; Mismatches 310; Indels 9; Gaps 3;

QY 24 GTTGATTTTCTCATTCACATATCCAAAGAAAACGTCGAGGGTCCCAAGATAAGTTTGTGAG 83  
Db 1 GCTTATTTCTCTACCCCTCCAAAGAAAAGAACTGAGGATCCCTTGATAAATTCGTTAA 60  
QY 84 TAAATTTCAAGGATGAGCTTTCCGTTGCTGCTTCACATCCAGTACCTGATGCTGCTCATCTT 143  
Db 61 GAAATTCAGGATGAACTTTTGTGCTGCCGCTCTCACTGTCTCTGACGGTCATGTTG 120  
QY 144 GCGTGTAGGACTAAGAGAAAGCTGACAAACAAAATTTGGTTTCAAGATGGTTGGCAAGATT 203  
Db 121 GCGTGTAGGAATAAGAAAGTTGACAAAGGTTTGGTTTTCAGGAAGTTTGGCAGAGTT 180  
QY 204 TGTGACCGTTTACTCCATTCGCAATTTGGTTATCTTTTGTATTTTGTATGATGAAAGAAATC 263  
Db 181 TCTAGAGCGTTACTATATTCGTTGCTTACGTAATTTGTTTGTATGATGAAAGAAATTC 240  
QY 264 TGCCTTCAGCGTCTACATTTTCAATTTATCCCACTCTGAGATCAATTACCATTCACCGG 323  
Db 241 TGCCTTCAGTGTAGTATATTTTAAATTTGTACAACTCGGAATTAATACTATCAGACTAATGC 300  
QY 324 TCTCATGATTCGCTCACAACCACTTCAAGCGCGCCGTTTGGTTTGAAGACCTTCAAGA 383  
Db 301 CCGTGTGTTACTCAATCAATACAGGAACAATATCCATTTGAAACAATTTGAAGATGA 360  
QY 384 TGAAGATGCCAGGTCATCTTTCTTCTGTGTACCATCACCACCTTCTCTGAGTCTAC 443  
Db 361 TGAATGCATCTCTCCAGCACTTTCAGAAATTTGTTTGGTGGTCTAAACTTAAACAATGCAT 420  
QY 444 AGTACAGCAACAAAGGATGCT---AGTTACGCAATCCAAACCTTTCTCCTGGACC 500  
Db 421 AAATCGGGTGTGTGTCGCAATCTTCAAACTTTCAAAAGGTTGTAATAAATCAACCTATTCCG 480  
QY 501 AGTTAAAGCTGAAGAGCCAAACGCCCAACCCCAAAATATACCTAA---AAAGAGAGGGAGAA 557  
Db 481 AGTCAAAATTCATATCTTCAAGTTTCAAGTGCATGAAACCAAGAACCAAAAAGCGTGGAG 540  
QY 558 GAAGAAAATGCTGATCTCTGAGAAATAAATCATCATCAGCTCCGCGAGATGATGATCCAGA 617  
Db 541 AAAGCGGAGTTTGTATCTTAACTGTCAGGATTCATCTGTCGACGTTGAAGATGATGTTGA 600  
QY 618 GAAACGTTCAAAGTTCTAGAGAGTCTTCTGCGAGAAAGAGAACCGTGAATGCGAGAGA 677  
Db 601 TATGCGCTTTAGATGTTATGAAAGTCTTTCAGCCCGGAAAGAGAACCGTGCAGACTGAAGA 660  
QY 678 AAGAGAGAGGCCATCAATGAGCCAAACCGTTTCGAAACCAACCAACCCCTTCTTCAGACT 737  
Db 661 AAGAGAGAGCAATCAATGAGCCAAACCGTTTCGAGCAATTAACCCCTTCTTCGAGGGT 720  
QY 738 GGTTCCTGCGACCATCTATCTATACAGAGTTGTCATGATCTTCTCTCTCTGCGGTTTGC 797  
Db 721 CGTCTTGGACCATCTTATCTGTACAGGGGATGATATGTAATCTTCTCTCTCTGCTTTCG 780  
QY 798 TGAGAAAGTACCTTAAGTGGGATCTCCGGGTTTCATCAAGTCCAGTTCGCGA---GAAACA 854  
Db 781 TGAGAAAGCATCTTAAGTGGGTTTCTGGATCCATTAATTAATTCAGCTTCTCTGATGGGAGACA 840  
QY 855 ATGGCTGTTTGGATGCTCTCAAAAGCCGGGAGAGCCAAATTCAGTCAAGGATGTTACGA 914  
Db 841 GTGCTCTGTAAGATGCTCTTTTAAAGAGGAGGCAAGGCTAATTTCACTGAGGGATGTTATGA 900  
QY 915 ATTAC 920  
Db 901 ATTAC 906

RESULT 15

CD813712

LOCUS CD813712 612 bp mRNA linear EST 10-JUL-2003  
DEFINITION BN15.020L13F020211 BN15 Brassica napus cDNA clone BN15020L13, mRNA  
sequence.  
ACCESSION CD813712  
VERSION CD813712.1 GI:32495652  
KEYWORDS EST.  
SOURCE Brassica napus (rape)  
ORGANISM Brassica napus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 612)  
Genoplante.

AUTHORS Genoplante, a major partnership french program in plant genomics

TITLE Genoplante, unpublished (2003)

JOURNAL Contact: Genoplante

COMMENT Genoplante  
93, rue Henri Rochefort 91025 EVRY CEDEX France  
Tel: 33 1 69 47 54 00  
Fax: 33 1 69 47 54 10

This sequence has been generated in the framework of the french  
plant genomics programme 'Genoplante' (<http://www.genoplante.com>  
and <http://genoplante-info.infobiogen.fr>).

FEATURES  
source  
1..612  
Location/Qualifiers  
/organism="Brassica napus"  
/mol\_type="mRNA"  
/cultivar="Jet Neuf"  
/db\_xref="taxon:3708"  
/clone="BN15020L13"  
/tissue\_type="seed"  
/clone\_lib="BN15"

## ORIGIN

Query Match 34.1%; Score 349.8; DB 6; Length 612;  
Best Local Similarity 90.6%; Pred. No. 1.2e-93;  
Matches 385; Conservative 0; Mismatches 37; Indels 3; Gaps 1;  
  
QY 1 ATGCCACGCCCTTTCTCCATAAGTTGATTTTCTCATCCACTATCCAGAAAAAGCTCTG 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
188 ATGCCACGCCCTTTCTCCACAGCTGATTTTCTCATCCACTATCCAGAAAAAGCTCTG 247  
QY 61 AGGGTCCCAAGTAAGTTGTGAGTAAATTCAGGATGAGCTTTCGGTTGCTGTGCACCTC 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
248 AGAGTTCCTGATAAGTTTGTGAGTAAATTCAGGACGAGCTATCGTTGCGGTGCGCTC 307  
QY 121 ACAGTACCTGATGGTCATGTTGGCGGTGAGGACTAAGGAAAGCTG---ACAACTAAAT 177  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
308 ACAGTACCTGATGGTCATGTTGGCGGTGAGGACTAAGGAAAGCTGACAACTAAAGATC 367  
QY 178 TGGTTTCAAGATGGTTGGCAAGATTTGTTGACCGTTACTCCATTCCGATTGGTTATCTT 237  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
368 TGGTTTCAAGATGGTTGGCAAGATTTGTTGACCGTTACTCAATCCGATTGGTTACCTT 427  
QY 238 TTGATTTTATGATATGAGAACTCTGCTTCAGCGTCTACATTTTCAATTTATCCAC 297  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
428 TTGATTTTATGATATGAGAACTCTGCTTCAGCGTCTACATTTTCACTTATCACAC 487  
QY 298 TCTGATCAATATACATTCACCGGTCTCATGATTCGCTTCACAAACACTTCAACGC 357  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
488 TCCGAGATCAATACACTCCACCGGTCTCATGATTCGACACAACTTCAACGCT 547  
QY 358 GCCCGTTGTTGAGACCTTGAAGATGAAGTCCGAGGTCTATCTTTCCTCTCTGTG 417  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
548 GCGCGTTGTTGAGACCTCGAAGATGAAGATGTTGAGGTGTTACCCCTTCTCTCTG 607  
QY 418 TACCC 422  
Db |||||  
608 TACCC 612

Search completed: December 30, 2004, 14:32:03  
Job time : 3907 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 13:02:12 ; Search time 98 Seconds  
(without alignments)  
2002.069 Million cell updates/sec

Title: US-10-088-187A-11  
Perfect score: 1796  
Sequence: 1 MRPFPFHKLIFSSTIQEKRL.....LLRTRDFVLKVTAFRVNEYV 341

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_02.\*

1: uniprot\_spport.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1796	100.0	341	2 Q813W1	Q813W1 arabidopsis
2	1513	84.2	329	2 Q6R2U8	Q6R2U8 brassica ca
3	1513	84.2	329	2 AAR92199	AAR92199 brassica
4	1509	84.0	329	2 Q6V0J6	Q6V0J6 brassica ca
5	1509	84.0	329	2 AAQ55453	AAQ55453 brassica
6	1197	66.6	230	2 Q9LJ67	Q9LJ67 arabidopsis
7	946	52.7	226	2 Q9XIB5	Q9XIB5 arabidopsis
8	942	52.4	226	2 Q94B43	Q94B43 arabidopsis
9	278	15.5	402	2 Q8S2E6	Q8S2E6 oryza sativ
10	251.5	14.0	750	2 Q851V0	Q851V0 oryza sativ
11	246	13.7	1029	2 Q851V5	Q851V5 oryza sativ
12	239	13.3	337	2 Q8RYD1	Q8RYD1 arabidopsis
13	229	12.8	190	2 Q9ZSH7	Q9ZSH7 arabidopsis
14	229	12.8	190	2 AAQ89626	AAQ89626 arabidops
15	227.5	12.7	461	2 Q9SZA5	Q9SZA5 arabidopsis
16	219.5	12.2	193	2 Q9LJ70	Q9LJ70 arabidopsis
17	215.5	12.0	209	2 Q84R27	Q84R27 arabidopsis
18	214	11.9	306	2 Q851V1	Q851V1 oryza sativ
19	203	11.3	307	2 Q9FUG2	Q9FUG2 arabidopsis
20	197	11.0	519	2 Q851W5	Q851W5 oryza sativ
21	195.5	10.9	301	2 Q9FK61	Q9FK61 arabidopsis
22	192	10.7	334	2 Q9FGD2	Q9FGD2 arabidopsis
23	191	10.6	427	2 Q7XS75	Q7XS75 oryza sativ
24	182.5	10.2	308	2 Q9LXB1	Q9LXB1 arabidopsis
25	179.5	10.0	302	2 Q9LSP6	Q9LSP6 arabidopsis
26	177.5	9.9	291	2 Q84RE8	Q84RE8 arabidopsis
27	175.5	9.8	545	2 Q851W4	Q851W4 oryza sativ
28	172	9.6	286	2 Q9XIG9	Q9XIG9 arabidopsis
29	172	9.6	389	2 Q9SZ05	Q9SZ05 arabidopsis
30	170.5	9.5	391	2 Q6Z0D2	Q6Z0D2 oryza sativ
31	170.5	9.5	391	2 BAD03253	BAD03253 oryza sat

RESULT 1

Q8L3W1	PRELIMINARY;	PRT;	341 AA.
ID	Q8L3W1;		
DT	01-OCT-2002 (TRENBLrel. 22, Created)		
DT	01-OCT-2002 (TRENBLrel. 22, Last sequence update)		
DT	05-JUL-2004 (TRENBLrel. 27, Last annotation update)		
DE	Reduced vernalization response 1.		
GN	Name=VRN1;		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsia.		
OX	NCBI_TaxID=3702;		
RN	[1]_TaxID=3702;		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22111275; PubMed=12114624;		
RA	Levy Y.Y., Meenage S., Mylne J.S., Gendall A.R., Dean C.;		
RT	"Multiple roles of Arabidopsis VRN1 in vernalization and flowering		
RT	time control."		
RL	Science 297:243-246(2002).		
DR	EMBL; AF289051; AAM76972.1; --		
DR	EMBL; AF289052; AAM76973.1; --		
DR	GO; GO:0003677; P:DNA binding; IEA.		
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.		
DR	InterPro; IPR003340; TF_B3.		
DR	Pfam; PF02362; B3; 2.		
DR	PROSITE; PS50863; B3; 2.		
SQ	SEQUENCE 341 AA; 39274 MW; E40ABFB2C8CC3114 CRC64;		
Query Match	100.0%;	Score 1796;	DB 2; Length 341;
Best Local Similarity	100.0%;	Pred. No. 2.1e-129;	
Matches 341;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1 MRPFPFHKLIFSSTIQEKRLRPDPKFKPDELSSVAVALTPDGHVVRVGLRKADNKIW 60		
Db	1 MRPFPFHKLIFSSTIQEKRLRPDPKFKPDELSSVAVALTPDGHVVRVGLRKADNKIW 60		
QY	61 FQDQGEFVDYRISIRIGYLLIFRYEGNSAFSVYIFNLHSEINHYSTGLMDSAHNHFKRA 120		
Db	61 FQDQGEFVDYRISIRIGYLLIFRYEGNSAFSVYIFNLHSEINHYSTGLMDSAHNHFKRA 120		
QY	121 RLFDLEDEDAEVIFFPSSVYPSPLPESTVPANKGYASSAIQTLFTGPVKAEEPTTPPKIP 180		
Db	121 RLFDLEDEDAEVIFFPSSVYPSPLPESTVPANKGYASSAIQTLFTGPVKAEEPTTPPKIP 180		
QY	181 KRGKKNKADPEEINNSAPRDDDDPENRSKFYESASAKRTVTABERERAINAAKTPEPT 240		
Db	181 KRGKKNKADPEEINNSAPRDDDDPENRSKFYESASAKRTVTABERERAINAAKTPEPT 240		
QY	241 NPFFRVLRPSLYRGCIIMYLPSPGPAEKYLSIGSFIKVLAEKQMPVRCLYKAGRAKFS 300		
Db	241 NPFFRVLRPSLYRGCIIMYLPSPGPAEKYLSIGSFIKVLAEKQMPVRCLYKAGRAKFS 300		

ALIGNMENTS

```

QY 301 QGWYFTLENNLGEQDVCVFELLRTDFVLKVTAPRVNEYV 341
DB 301 QGWYFTLENNLGEQDVCVFELLRTDFVLKVTAPRVNEYV 341

RESULT 2
Q6R2U8 PRELIMINARY; PRT; 329 AA.
AC Q6R2U8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Reduced vernalization response 1.
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]_
RP SEQUENCE FROM N.A.
RA Kwon S.-J., Park B.-S., Kim S.-Y., Choi H.-S., Lee M.-C., Kim J.-S.,
RA Lee S.-I., Lim K.-B., Kim J.-A., Lee M.-R., Jin Y.-M., Kim D.,
RA Kim H.-I.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY517929; AAR92199.1; -.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 2.
DR PROSITE; PS50863; B3; 2.
SQ SEQUENCE 329 AA; 38058 MW; A24ADB91433459BB CRC64;

Query Match 84.2%; Score 1513; DB 2; Length 329;
Best Local Similarity 85.1%; Pred. No. 1e-107;
Matches 296; Conservative 13; Mismatches 13; Indels 26; Gaps 5;

QY 1 MPRPFHKLIFSSTIQEKLRLVPDKFVSKFDLSVAVALTPDGHVWRVGLRKAD--NH 58
DB 1 MPRPFHKLIFSSTIQEKLRLVPDKFVSKFDLSVAVALTPDGHVWRVGLRKADNNK 60

QY 59 IFQDQWQEFVDYRYSIRIGYLLIFRYEGNSAFSVYIFNLSHSEINHYSTGLMDSA--HNH 116
DB 61 IFQDQWQEFVDYRYSIRIGYLLIFRYEGNSAFSVYIFNLSHSEINHYSTGLMDSASHNNH 120

QY 117 FKARLFEDELEDAEVIFFSPVPSPLPSTVPANKGYASSAIQTUFTGPVKAEE--PT 174
DB 121 FKARLFEDELEDAETLHTT-----ASAIQSFTGPKPEEATPT 161

QY 175 PTPKPKKGRKKKNAD-PEEINSSAPRDDDDPENRSKFYESASARKRTVTAEERERAINA 233
DB 162 QTSKVPKKGKKNADHPPEVNSSAPRDDDDPENRSKFYESASARKRTVTAEERERAVNA 221

QY 234 AKTPEPTNPFFRVLRPSLYRGCIWYLPSPGFAEKYLSGISGFIKVLAEKQWPVRCLYK 293
DB 222 AKTPEPTNPFFRVLRPSLYRGCIWYLPSPGFAEKYLSGISGFIKVLAEKQWPVRCLYK 281

QY 294 AGRAKFSQGWYFTLENNLGEQDVCVFELLRTDFVLKVTAPRVNEYV 341
DB 282 AGRAKFSQGWYFTVENNLGEQDVCVFELLRTDFVLKVTAPRVNEYV 329

RESULT 4
Q6V0J6 PRELIMINARY; PRT; 329 AA.
AC Q6V0J6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Reduced vernalization response 1.
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]_
RP SEQUENCE FROM N.A.
RA Kwon S.-J., Park B.-S., Kim S.-Y., Choi H.-S., Lee M.-C., Kim J.-S.,
RA Lee S.-I., Lim K.-B., Kim J.-A., Lee M.-R., Jin Y.-M.,
RA Kim D., Kim H.-I.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY356368; AAQ55453.1; -.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 2.
DR PROSITE; PS50863; B3; 2.
SQ SEQUENCE 329 AA; 37931 MW; BFF71D55ADE2655A CRC64;

Query Match 84.0%; Score 1509; DB 2; Length 329;
Best Local Similarity 85.1%; Pred. No. 2e-107;
Matches 296; Conservative 12; Mismatches 14; Indels 26; Gaps 5;

QY 1 MPRPFHKLIFSSTIQEKLRLVPDKFVSKFDLSVAVALTPDGHVWRVGLRKAD--NH 58
DB 1 MPRPFHKLIFSSTIQEKLRLVPDKFVSKFDLSVAVALTPDGHVWRVGLRKADNNK 60

QY 59 IFQDQWQEFVDYRYSIRIGYLLIFRYEGNSAFSVYIFNLSHSEINHYSTGLMDSA--HNH 116

```

```

QY 301 QGWYFTLENNLGEQDVCVFELLRTDFVLKVTAPRVNEYV 341
DB 301 QGWYFTLENNLGEQDVCVFELLRTDFVLKVTAPRVNEYV 341

RESULT 2
Q6R2U8 PRELIMINARY; PRT; 329 AA.
AC Q6R2U8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Reduced vernalization response 1.
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]_
RP SEQUENCE FROM N.A.
RA Kwon S.-J., Park B.-S., Kim S.-Y., Choi H.-S., Lee M.-C., Kim J.-S.,
RA Lee S.-I., Lim K.-B., Kim J.-A., Lee M.-R., Jin Y.-M., Kim D.,
RA Kim H.-I.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY517929; AAR92199.1; -.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 2.
DR PROSITE; PS50863; B3; 2.
SQ SEQUENCE 329 AA; 38058 MW; A24ADB91433459BB CRC64;

Query Match 84.2%; Score 1513; DB 2; Length 329;
Best Local Similarity 85.1%; Pred. No. 1e-107;
Matches 296; Conservative 13; Mismatches 13; Indels 26; Gaps 5;

QY 1 MPRPFHKLIFSSTIQEKLRLVPDKFVSKFDLSVAVALTPDGHVWRVGLRKAD--NH 58
DB 1 MPRPFHKLIFSSTIQEKLRLVPDKFVSKFDLSVAVALTPDGHVWRVGLRKADNNK 60

QY 59 IFQDQWQEFVDYRYSIRIGYLLIFRYEGNSAFSVYIFNLSHSEINHYSTGLMDSA--HNH 116
DB 61 IFQDQWQEFVDYRYSIRIGYLLIFRYEGNSAFSVYIFNLSHSEINHYSTGLMDSASHNNH 120

QY 117 FKARLFEDELEDAEVIFFSPVPSPLPSTVPANKGYASSAIQTUFTGPVKAEE--PT 174
DB 121 FKARLFEDELEDAETLHTT-----ASAIQSFTGPKPEEATPT 161

QY 175 PTPKPKKGRKKKNAD-PEEINSSAPRDDDDPENRSKFYESASARKRTVTAEERERAINA 233
DB 162 QTSKVPKKGKKNADHPPEVNSSAPRDDDDPENRSKFYESASARKRTVTAEERERAVNA 221

QY 234 AKTPEPTNPFFRVLRPSLYRGCIWYLPSPGFAEKYLSGISGFIKVLAEKQWPVRCLYK 293
DB 222 AKTPEPTNPFFRVLRPSLYRGCIWYLPSPGFAEKYLSGISGFIKVLAEKQWPVRCLYK 281

QY 294 AGRAKFSQGWYFTLENNLGEQDVCVFELLRTDFVLKVTAPRVNEYV 341
DB 282 AGRAKFSQGWYFTVENNLGEQDVCVFELLRTDFVLKVTAPRVNEYV 329

RESULT 3
AAR92199 PRELIMINARY; PRT; 329 AA.
AC AAR92199;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Reduced vernalization response 1.
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]_

```

```
Db 61 IMFQDQWQFVDRYSIRIGYLLIFRYEGNSAFSVCIYNLPQSEINHYSTGLMDSASHNNH 120
QY 117 FKARLFEDLEDEDAEVIFFPSVYPSLPSTVPANKGVASSAIOFLFTGPVKAER--PT 174
Db 121 FKRPFLFEDLEDEDAETLHTT-----ASAIQSFTFGPVKPEATPT 161
QY 175 PTPKIPKGRKKKNAD-PEINSSAPRDDDDPENRSKPFYESASARKRTVTAEERERAINA 233
Db 162 QTSKVPKGRKKKNADHPEEVNSSAPRDDDDPENRSKPFYESASARKRTVTAEERERAVNA 221
QY 234 AKTEFTNPFFRVLRPSLYRGCMYLPSPGAEKYLSIGSFIVQLAEKQWVRCLYK 293
Db 222 AKTFEFTNPFFRVLRPSLYRGCMYLPSPGAEKYLSIGSFIVQLAEKQWVRCLYK 281
QY 294 AGRAKFSQGWYEFTELENNLGEQDVCVFELLRTDRDFVLKVTAPRVNEYV 341
Db 282 AGRAKFSQGWYEFTELENNLGEQDVCVFELLRTDRDFVLKVTAPRVNEYV 329

RESULT 5
AAQ55453
ID AAQ55453 PRELIMINARY; PRT; 329 AA.
AC AAQ55453;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE Reduced vernalization response 1.
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cev. Samjin;
RA Kwon S.-J., Park B.-S., Kim S.-Y., Choi H.-S., Lee M.-C., Kim J.-S.,
RA Lee S.-I., Lim K.-B., Kim J.-A., Hong K.-Y., Lee M.-R., Jin Y.-M.,
RA Kim D., Kim H.-I.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY356368; AAQ55453.1; -- BFF71D55ADE2655A CRC64;
SQ SEQUENCE 329 AA; 37931 MW; BFF71D55ADE2655A CRC64;

Query Match 84.0%; Score 1509; DB 2; Length 329;
Best Local Similarity 85.1%; Pred. No. 2e-107;
Matches 296; Conservative 12; Mismatches 14; Indels 26; Gaps 5;

QY 1 MRPFPFHKLIFSSTIQEKLRLVDPKFSKFDLSVAVALTPDGHVWRVGLRKAD--NK 58
Db 1 MRPFPFHKLIFSSTIQEKLRLVDPKFSKFDLSVAVALTPDGHVWRVGLRKADNNK 60
QY 59 IMFQDQWQFVDRYSIRIGYLLIFRYEGNSAFSVCIYNLPQSEINHYSTGLMDSA--HNH 116
Db 61 IMFQDQWQFVDRYSIRIGYLLIFRYEGNSAFSVCIYNLPQSEINHYSTGLMDSASHNNH 120
QY 117 FKARLFEDLEDEDAEVIFFPSVYPSLPSTVPANKGVASSAIOFLFTGPVKAER--PT 174
Db 121 FKRPFLFEDLEDEDAETLHTT-----ASAIQSFTFGPVKPEATPT 161
QY 175 PTPKIPKGRKKKNAD-PEINSSAPRDDDDPENRSKPFYESASARKRTVTAEERERAINA 233
Db 162 QTSKVPKGRKKKNADHPEEVNSSAPRDDDDPENRSKPFYESASARKRTVTAEERERAVNA 221
QY 234 AKTEFTNPFFRVLRPSLYRGCMYLPSPGAEKYLSIGSFIVQLAEKQWVRCLYK 293
Db 222 AKTFEFTNPFFRVLRPSLYRGCMYLPSPGAEKYLSIGSFIVQLAEKQWVRCLYK 281
QY 294 AGRAKFSQGWYEFTELENNLGEQDVCVFELLRTDRDFVLKVTAPRVNEYV 341
Db 282 AGRAKFSQGWYEFTELENNLGEQDVCVFELLRTDRDFVLKVTAPRVNEYV 329

RESULT 6
```

```
Q9LJ67
ID Q9LJ67 PRELIMINARY; PRT; 230 AA.
AC Q9LJ67;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Gb/AAD43153.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety pl,
RT TAC and BAC clones.";
RL DNA Res. 7:217-221(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP000735; BAB01695.1; --
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 1.
DR PROSITE; PS50863; B3; 1.
SQ SEQUENCE 230 AA; 26219 MW; B96A826B384C9F3C CRC64;

Query Match 66.6%; Score 1197; DB 2; Length 230;
Best Local Similarity 99.1%; Pred. No. 1e-83;
Matches 230; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 110 MDSAHNHFKARLFEDLEDEDAEVIFFPSVYPSLPSTVPANKGVASSAIOFLFTGPVK 169
Db 1 MDSAHNHFKARLFEDLEDEDAEVIFFPSVYPSLPSTVPANKGVASSAIOFLFTGPVK 60
QY 170 ABEPTTPKIPKGRKKKNADPEINSSAPRDDDDPENRSKPFYESASARKRTVTAEERER 229
Db 61 --EPTTPKIPKGRKKKNADPEINSSAPRDDDDPENRSKPFYESASARKRTVTAEERER 118
QY 230 AINAAKTEPTNPFFRVLRPSLYRGCMYLPSPGAEKYLSIGSFIVQLAEKQWVR 289
Db 119 AINAAKTEPTNPFFRVLRPSLYRGCMYLPSPGAEKYLSIGSFIVQLAEKQWVR 178
QY 290 CLYKAGRAKFSQGWYEFTELENNLGEQDVCVFELLRTDRDFVLKVTAPRVNEYV 341
Db 179 CLYKAGRAKFSQGWYEFTELENNLGEQDVCVFELLRTDRDFVLKVTAPRVNEYV 230

RESULT 7
Q9XIB5
ID Q9XIB5 PRELIMINARY; PRT; 226 AA.
AC Q9XIB5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F13F21.8 protein.
GN Name=F13F21.8;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federpspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
```

```

RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007504; AAD3153.1; -.
DR PIR; E96531; E96531.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 1.
DR PROSITE; PS50863; B3; 1.
SQ SEQUENCE 226 AA; 25850 MW; CB902C811E446A67 CRC64;

Query Match 52.7%; Score 946; DB 2; Length 226;
Best Local Similarity 81.2%; Pred. No. 1.7e-64;
Matches 190; Conservative 14; Mismatches 18; Indels 12; Gaps 5;

QY 110 MDSAHNF-KRRLPFEDLEDAEVIFFSSVYPSLPSTVPANKGY-ASSAIQTLFTGP 167
Db GO:0003677; F:DNA binding; IEA.
3 MDSAQNFNKRARLPEDPELKDVKIYPSN-----PESTFPVNGYGGSTAIQSFYK-E 55
QY 168 VKAEETPTPKIPKGRKKKNADPEEINSSAPRDDDPENRSKFVESASARKRTVTAEER 227
Db GO:0003677; F:DNA binding; IEA.
56 SKAEE---TPKVLKGRKKKNPNPEEVSNSSTPGDDSENRSKFVESASARKRTVTAEER 112
QY 228 ERAINAAKTFFETNPFVRVLRPSYLRGCMYLPSPGFAEKYLSGIFIKVQLAEKQWP 287
Db GO:0003677; F:DNA binding; IEA.
113 ERAVNAKTFFETNPFVRVLRPSYLRGCMYLPSPGFAEKYLSGIFIKVQLAEKQWP 172
QY 288 VRCLYKAGRAKFSQGWYEFTELENNIGEGDVCVFELLRTDRFVLKVTAFRNEYV 341
Db GO:0003677; F:DNA binding; IEA.
173 VRCLYKAGRAKFSQGWYEFTELENNIGEGDVCVFELLRTDRFVLKVTAFRNEYV 226

Query Match 52.4%; Score 942; DB 2; Length 226;
Best Local Similarity 80.8%; Pred. No. 3.5e-64;
Matches 189; Conservative 13; Mismatches 20; Indels 12; Gaps 5;

QY 110 MDSAHNF-KRRLPFEDLEDAEVIFFSSVYPSLPSTVPANKGYASSA-IQTLFTGP 167
Db GO:0003677; F:DNA binding; IEA.
3 MDSAQNFNKRARLPEDPELKDVKIYPSN-----IPSTFPVNGYGGSTDIQSFYK-E 55
QY 168 VKAEETPTPKIPKGRKKKNADPEEINSSAPRDDDPENRSKFVESASARKRTVTAEER 227
Db GO:0003677; F:DNA binding; IEA.

Query Match 52.4%; Score 278; DB 2; Length 402;
Best Local Similarity 26.8%; Pred. No. 5.9e-13;
Matches 103; Conservative 57; Mismatches 157; Indels 68; Gaps 16;

QY 3 RPFFKHLIFSTIQEKLRLVPDKFV-----SKFKDELVAVALTPDGHVWRVGL 52
Db RPFFKHLIFSTIQEKLRLVPDKFV-----SKFKDELVAVALTPDGHVWRVGL 52
QY 18 RPHFFKVLVGDFF--KQRLKIPNFCKHLPWEESKAKGLKEASNAATLEGSGGTWLVI 75
Db RPHFFKVLVGDFF--KQRLKIPNFCKHLPWEESKAKGLKEASNAATLEGSGGTWLVI 75
QY 53 RKADNKIWFQDQWQBFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLHSEINHYHSTGLMDS 112
Db RKADNKIWFQDQWQBFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLHSEINHYHSTGLMDS 112
QY 76 RRTAEGIFFTSGWPKFVDQALRELFVWFVRYDGNTRFTAMVFRTACE----REDLMGG 131
Db RRTAEGIFFTSGWPKFVDQALRELFVWFVRYDGNTRFTAMVFRTACE----REDLMGG 131
QY 113 A---HHNFKARLPEDLEDA-----EVI-----FPSSVYP-----SPLPES 147
Db A---HHNFKARLPEDLEDA-----EVI-----FPSSVYP-----SPLPES 147
QY 132 GGGDRPRKGRGRPTAAASRDARPKKDSVGKEMVTVRASFGSQGLQIVDSWTPPEGS 191
Db GGGDRPRKGRGRPTAAASRDARPKKDSVGKEMVTVRASFGSQGLQIVDSWTPPEGS 191
QY 148 TVPANKGYASSAIQTLFTGPVKAEEPTPKIPK-----KRGKKKNADPEEIN- 196
Db TVPANKGYASSAIQTLFTGPVKAEEPTPKIPK-----KRGKKKNADPEEIN- 196
QY 192 TAVKNEEDADE---LPCVCELPASSASFPFVHPGALDADGGAARRGAAKRTSLQDLDAL 247
Db TAVKNEEDADE---LPCVCELPASSASFPFVHPGALDADGGAARRGAAKRTSLQDLDAL 247

RESULT 8
Q94B43 ID Q94B43 PRELIMINARY; PRT; 226 AA.
AC Q94B43;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein F13F21.8.
GN Name=F13F21.8;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayaehizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shimizu P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY042868; AAK68808.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 1.
DR PROSITE; PS50863; B3; 1.
KW Hypothetical protein.
SQ SEQUENCE 226 AA; 25893 MW; 2093F693B98CB075 CRC64;

Query Match 52.4%; Score 942; DB 2; Length 226;
Best Local Similarity 80.8%; Pred. No. 3.5e-64;
Matches 189; Conservative 13; Mismatches 20; Indels 12; Gaps 5;

QY 110 MDSAHNF-KRRLPFEDLEDAEVIFFSSVYPSLPSTVPANKGYASSA-IQTLFTGP 167
Db GO:0003677; F:DNA binding; IEA.
3 MDSAQNFNKRARLPEDPELKDVKIYPSN-----IPSTFPVNGYGGSTDIQSFYK-E 55
QY 168 VKAEETPTPKIPKGRKKKNADPEEINSSAPRDDDPENRSKFVESASARKRTVTAEER 227
Db GO:0003677; F:DNA binding; IEA.

Query Match 52.7%; Score 946; DB 2; Length 226;
Best Local Similarity 81.2%; Pred. No. 1.7e-64;
Matches 190; Conservative 14; Mismatches 18; Indels 12; Gaps 5;

QY 110 MDSAHNF-KRRLPFEDLEDAEVIFFSSVYPSLPSTVPANKGY-ASSAIQTLFTGP 167
Db GO:0003677; F:DNA binding; IEA.
3 MDSAQNFNKRARLPEDPELKDVKIYPSN-----PESTFPVNGYGGSTAIQSFYK-E 55
QY 168 VKAEETPTPKIPKGRKKKNADPEEINSSAPRDDDPENRSKFVESASARKRTVTAEER 227
Db GO:0003677; F:DNA binding; IEA.
56 SKAEE---TPKVLKGRKKKNPNPEEVSNSSTPGDDSENRSKFVESASARKRTVTAEER 112
QY 228 ERAINAAKTFFETNPFVRVLRPSYLRGCMYLPSPGFAEKYLSGIFIKVQLAEKQWP 287
Db GO:0003677; F:DNA binding; IEA.
113 ERAVNAKTFFETNPFVRVLRPSYLRGCMYLPSPGFAEKYLSGIFIKVQLAEKQWP 172
QY 288 VRCLYKAGRAKFSQGWYEFTELENNIGEGDVCVFELLRTDRFVLKVTAFRNEYV 341
Db GO:0003677; F:DNA binding; IEA.
173 VRCLYKAGRAKFSQGWYEFTELENNIGEGDVCVFELLRTDRFVLKVTAFRNEYV 226

RESULT 9
Q8S2E6 ID Q8S2E6 PRELIMINARY; PRT; 402 AA.
AC Q8S2E6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE P0022F10.13 protein.
GN Name=P0022F10.13;
OS Eukarya sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Niumura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakanishi Y., Nakamura M.,
RA Nami N., Negishi M., Ohta I., Ono N., Sai S., Sakai, K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Teuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojibori T.;
RL "The genome sequence and structure of rice chromosome 1."
RT Nature 420:312-316(2002).
DR EMBL; AP003229; BAB89497.1; -.
DR Gramene; Q8S2E6; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 2.
DR PROSITE; PS50863; B3; 2.
SQ SEQUENCE 402 AA; 44783 MW; CFCALF234BC35F76 CRC64;

Query Match 15.5%; Score 278; DB 2; Length 402;
Best Local Similarity 26.8%; Pred. No. 5.9e-13;
Matches 103; Conservative 57; Mismatches 157; Indels 68; Gaps 16;

QY 3 RPFFKHLIFSTIQEKLRLVPDKFV-----SKFKDELVAVALTPDGHVWRVGL 52
Db RPFFKHLIFSTIQEKLRLVPDKFV-----SKFKDELVAVALTPDGHVWRVGL 52
QY 18 RPHFFKVLVGDFF--KQRLKIPNFCKHLPWEESKAKGLKEASNAATLEGSGGTWLVI 75
Db RPHFFKVLVGDFF--KQRLKIPNFCKHLPWEESKAKGLKEASNAATLEGSGGTWLVI 75
QY 53 RKADNKIWFQDQWQBFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLHSEINHYHSTGLMDS 112
Db RKADNKIWFQDQWQBFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLHSEINHYHSTGLMDS 112
QY 76 RRTAEGIFFTSGWPKFVDQALRELFVWFVRYDGNTRFTAMVFRTACE----REDLMGG 131
Db RRTAEGIFFTSGWPKFVDQALRELFVWFVRYDGNTRFTAMVFRTACE----REDLMGG 131
QY 113 A---HHNFKARLPEDLEDA-----EVI-----FPSSVYP-----SPLPES 147
Db A---HHNFKARLPEDLEDA-----EVI-----FPSSVYP-----SPLPES 147
QY 132 GGGDRPRKGRGRPTAAASRDARPKKDSVGKEMVTVRASFGSQGLQIVDSWTPPEGS 191
Db GGGDRPRKGRGRPTAAASRDARPKKDSVGKEMVTVRASFGSQGLQIVDSWTPPEGS 191
QY 148 TVPANKGYASSAIQTLFTGPVKAEEPTPKIPK-----KRGKKKNADPEEIN- 196
Db TVPANKGYASSAIQTLFTGPVKAEEPTPKIPK-----KRGKKKNADPEEIN- 196
QY 192 TAVKNEEDADE---LPCVCELPASSASFPFVHPGALDADGGAARRGAAKRTSLQDLDAL 247
Db TAVKNEEDADE---LPCVCELPASSASFPFVHPGALDADGGAARRGAAKRTSLQDLDAL 247

```

```
QY 197 SSAPRDDPENRSKFYASARKTVAERERAINAAKTFTPTNFFRVLRPSVLYRG 256
DB 248 ASIP-----PSIRR--YKGVSRRAVATAERQATEIAHAFRSPLYCVMRSTMHVYYS 301
QY 257 CMIYLPSPGAEKYL-SGISGFIKVLQAEKQWVRCLYKAG-RAKFSQGWVEFTLENNLGE 314
DB 302 FMRFTGFSRQHLPRRTDVLVRDGGKWSV--LYIPNTRDLRSLRGWCAFGNCLSE 359
QY 315 GDVCPPELLRTRDFVLKVTAFRVNE 339
DB 360 GDVCPVELVAAREF--RVHIFRVE 382

RESULT 10
Q851V0
ID Q851V0 PRELIMINARY; PRT; 750 AA.
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DE Putative auxin response factor.
GN Name=OSJNB0111B07.22;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Ziemann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.O., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Ziemann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.O., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC097280; AAC34502.1; -.
DR Gramene; Q851V0; -.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 4.
DR PROSITE; PS0863; B3; 4.
SQ SEQUENCE 750 AA; 83502 MW; E9E3980EF8EF3E3 CRC64;

Query Match 14.0%; Score 251.5; DB 2; Length 750;
Best Local Similarity 26.9%; Pred. No. 1.5e-10;
Matches 86; Conservative 55; Mismatches 134; Indels 45; Gaps 13;

QY 22 VPDKFSVKFDELSVAVALTPDGHVVRVGLKAD-NK1WFDQWQEFVDRYSIRIGYLL 80
DB 424 VPARFANNFGHISEEVNLRSPSGETWSTGVANSAGELVLPQWKEFVGDNGISEGDCI 483
QY 81 IFRYEG-NSAFSVYIFNLSHSEINYSHTGLMDSAHNHFKRRLPEDLEDE-DAEVIFFSS 138
DB 484 LFRYGVSSSFVLLIPDPGCE-----KASPHFVSGHFGRAENSAGAEQGGNG 533
QY 139 VYPSPLPESTVPANKGYASSAQT1FTGPVKABEPTTPKIPK--KRG-----RKKN 189
DB 534 RTTPPI-----VDGNGRHHLENTLHNSCRS-----IPACKRS1FSDTEAKEND 581
QY 190 ADEEINSAPRDDPENRSKFYASAR--KRTVTAERERAINAAKTFTPTNFFRVV 247
DB 582 GEDEDVAAA-----EGGRYGEYFSGHGRVABYNLRREDEBISRVVPVQPCNPVFQV 637
QY 248 LPSVYL--YRGCMYLPSPGAEKYLSGISGFIKVLQAEK--OWPVRCLYKAG-RAKFSQ 302
DB 638 IHSSHVRSKYCIYGVSPFAGKYLGAVERVVLRAERSGGEMHVPFVRHQNTRFGYAG 697
```

```
QY 303 WYEFTLENNLSEGDVCFEL 322
DB 698 WRQFAGDNLRLVAHDVCLPEL 717

RESULT 11
Q851V5
ID Q851V5 PRELIMINARY; PRT; 1029 AA.
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Putative auxin response factor.
GN Name=OSJNB0111B07.17;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Ziemann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.O., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC097280; AAC34491.1; -.
DR Gramene; Q851V5; -.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003340; TF_B3.
DR Pfam; PF02362; B3; 5.
DR PROSITE; PS0863; B3; 5.
SQ SEQUENCE 1029 AA; 118189 MW; 206B6AD0CA4B7645 CRC64;

Query Match 13.7%; Score 246; DB 2; Length 1029;
Best Local Similarity 23.8%; Pred. No. 5.9e-10;
Matches 74; Conservative 60; Mismatches 123; Indels 54; Gaps 8;

QY 5 PFHKLIFESTIQEKELRVPDKFVSKFDELSVAVALTPDGHVVRVGLRKADNKIWFQDG 64
DB 149 YFFKVMIGGF--RQMTIPYKFAENFRDQIQGTIKLKARNGTCSVLVDKCSNKLVLTKG 206
QY 65 WQEFVDYRYSIRIGYLLIFRYEGNSAFSVYIFNLSHSEINYSHTGLMDSA-HNHFKRRLF 123
DB 207 WAEPFANSHDKMGDFLVRFTYTGNSQFVKIFDPS-----GCVKAASHNAVNIQHA 257
QY 124 EDLEDEDAEVIFFPSVYPSPLPESTVPANKGYASSAQT1FTGPVKABEPTTPKIPK 183
DB 258 QNMQGDPTIILSCSDEHLRAQSLTTERQNQ-----PEKD 291
QY 184 GRKKKNADPEEINSAPRDDPENRSKFYASARKTVAERERAINAAKTFTPTNPF 243
DB 292 VIDNCNKKKKTEHASSSEDQ-----ETPTAEVHRMKVEEMVRAIHS-----NHVP 337
QY 244 FRVLVRPSVLYR-GCMYLPSPGAEKYLSGISGFIKVLQAEKQWVR-CLYKAGRAKFSQ 301
DB 338 FVAVMKSNVTPQPCYVAISRKYANEYFPGDQMLTLQRHGRWQVKFCISKKLEMLSK 397
QY 302 GWYEFTLENNL 312
DB 398 GWRKFTRDNEL 408

RESULT 12
Q8RYD1
ID Q8RYD1 PRELIMINARY; PRT; 337 AA.
```

Stoneking T., Smith R.;  
Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
[3]  
SEQUENCE FROM N.A.  
Waterston R.;  
Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
[4]  
SEQUENCE FROM N.A.  
Cheuk R., Chen H.A., Kim C.J., Shinn P., Carninci P., Hayashizaki Y.,  
Ishida J., Kamiya A., Kawai J., Narusaka M., Sakurai T., Satou M.,  
Seki M., Shinozaki K., Ecker J.R.;  
Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.  
[5]  
SEQUENCE FROM N.A.  
Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,  
Mayer K.F.X.;  
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[6]  
SEQUENCE FROM N.A.  
EU Arabidopsis sequencing project;  
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
EMBL; AF104919; AAC72857.1; -;  
EMBL; BT010604; AAQ89626.1; -;  
EMBL; AL161492; CAB7728.1; -;  
PIR; T02015; T02015.  
GO; GO:0003677; F:DNA binding; IEA.  
GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
InterPro; IPR008972; Cypredoxin.  
InterPro; IPR003340; TF\_B3.  
Pfam; PF02362; B3; 1.  
PROSITE; PS0863; B3; 1.  
Hypothetical protein.  
SEQUENCE 190 AA; 22122 MW; D7FB2C343E4BC915 CRC64;  
SQ  
Query Match 12.8%; Score 229; DB 2; Length 190;  
Best Local Similarity 33.3%; Pred. No. 1.2e-09;  
Matches 53; Conservative 25; Mismatches 51; Indels 30; Gaps 4;  
QY 6 FHKLIPSSITQBRKRLRVPDKFYKFDKESVAVALTPDGHVWRVGLRKADNKIWFODGW 65  
DB 30 FFKVLVLPSTMDKMMRIPRFVKLQSGKLSVVTLTPAGYKRSIKLGRIGEIWPFHEG 89  
QY 66 QFVDVRYSTRIGVLLIFRFGHSAFVYIENLSHSEINY-----HSTGL- 109  
DB 90 SFBAEASHIEGHFLFLFYKKNSSFRVLIIFNACSTNVPLDVAHIIDSDDDVIEITGKE 149  
QY 110 MDSANHHFKRAR-----LFDL-----EDEDAEVI 134  
DB 150 FTEHKSCKRPRDIEFDKILHDVDVMQVLKEEEDKRVL 188  
RESULT 14  
AAQ89626 PRELIMINARY; PRT; 190 AA.  
ID AAQ89626  
AC AAQ89626;  
DT 02-MAR-2004 (TEMBLrel. 27, Created)  
DT 02-MAR-2004 (TEMBLrel. 27, Last sequence update)  
DT 02-MAR-2004 (TEMBLrel. 27, Last annotation update)  
DE At4g01580.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosid II; Brassicales; Brassicaceae; Arabidopsais.  
NCBI\_TaxID=3702;  
[1]  
SEQUENCE FROM N.A.  
Cheuk R., Chen H.A., Kim C.J., Shinn P., Carninci P., Hayashizaki Y.,  
Ishida J., Kamiya A., Kawai J., Narusaka M., Sakurai T., Satou M.,  
Seki M., Shinozaki K., Ecker J.R.;  
"Arabidopsis ORF clones".  
Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.  
EMBL; BT010604; AAQ89626.1; -;  
SQ SEQUENCE 190 AA; 22122 MW; D7FB2C343E4BC915 CRC64;

Query Match 12.8%; Score 229; DB 2; Length 190;  
Best Local Similarity 33.3%; Pred. No. 1.2e-09;  
Matches 53; Conservative 25; Mismatches 51; Indels 30; Gaps 4;  
QY 6 FHKLISSIOEKRLAVDPKFKVSKFDKLSVAVALTPDGHVWRVGLRKADNKNIFQDQW 65  
DB 30 FFKLVLPSTWKQKQRIIPRFVFKLQSGSKLSEVTVLTPAGYKRSIKLKRIGEINIFHEGW 89  
QY 66 QEFVDYSIRIGYLLIFRYEGNSAFSVYIFNLHSHSINY-----HSTGL- 109  
DB 90 SEFAEASHTEEGHFLFEYKNSFRVILFNASACSTNPLDAVHIDSDDDVIEITGKE 149  
QY 110 MOSAHNFKRAR-----LFEDL-----EDEDAEVI 134  
DB 150 FDEHKSKRPRDIEFDKILHDVDVQVLKEBEDKRVL 188

RESULT 15

Q9SZA5 PRELIMINARY; PRT; 461 AA.  
AC Q9SZA5;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein F17M5.40 (Hypothetical protein AT4g33280).  
GN Name=F17M5.40; Synonyms=AT4g33280;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bevan M., Rose M., Hempel S., Entian K.-D., Hoheisel J., Mewes H.W.,  
RA Mannhaupt G., Mayer K.F.X., Schueller C.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL035678; CAB38786.1; -;  
DR EMBL; AL161583; CAB80045.1; -;  
DR PIR; T05979; T05979.  
DR GO; GO:0003677; F.DNA binding; IEA.  
DR GO; GO:0006355; P.regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR003340; TF\_B3.  
DR Pfam; PF02362; B3; 2.  
DR PROSITE; PS00863; B3; 2.  
KW Hypothetical protein.  
SQ SEQUENCE 461 AA; 52997 MW; D7EB497600A33062 CRC64;

Query Match 12.7%; Score 227.5; DB 2; Length 461;  
Best Local Similarity 23.3%; Pred. No. 5.3e-09;  
Matches 86; Conservative 63; Mismatches 139; Indels 81; Gaps 13;  
QY 22 VPDKFKVSKDELSVAVALTPDGHVWRVGLRKADNK-IWFQDQWQEFVDRYSIRIGYLL 80  
DB 108 IPRKFTCKRKLQIVITLKSFGVTVNGVEEDDEKTAFRGWDKFKVXKHSLENDLL 167  
QY 81 IFRYEGNSAFSVYIFN-----LSHSE-----INYSHTGLMDSAHNHFKRA 120  
DB 168 VFKFGVSEFEVLVFDGQILCEKPTSYFVRKCGHAETKGIIDFNATS-SRSPKRFHND 226  
QY 121 RL-----FEDLEDEAEVIFPSSVYPSPLPESTVPANKGYASSAIQTL 163

Search completed: December 30, 2004, 06:10:40  
Job time : 100 secs

DB 227 DVETTPNQQLVISPPVDNELEDLIDIDLDIDKILNPLLVASHTGYEQEHHNSDIDA 286  
QY 164 FTGPPVKAEBPTPTPKIP-----KKGRKKKNADPEEINSSAPRDDDDPENRSKFYES 214  
DB 287 -SAQLPVISPTSTVRSSEKYPKSGFKGWRRLSNDNLQ-----KA 327  
QY 215 ASARKRTVTAERERAINAAKTFBPTNPPFVVLRLPSYLYRGCIWYLPSPGFAEKYLSGIS 274  
DB 328 AGSNKKALS-----AKRAISPDG--FLVFMKSHVVSKFCLTIPYKWCVKWMLIR 377  
QY 275 GFIKVQLAEKQWPVRC-LYKA-GRAKFSQGWTEFTLENNLGGDVCVPELL--RTDRFVL 330  
DB 378 QEVVMQVDQTKWEMKFNIFGARGSGGISTGKFKFQDNNLRGDDVCVFPANSETKPLHL 437  
QY 331 KVTAERVNE 339  
DB 438 NVYIFRGE 446

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 19:00:56 ; Search time 23 Seconds  
(without alignments)  
1426.518 Million cell updates/sec

Title: US-10-088-187A-11  
Perfect score: 1796  
Sequence: 1 MRPFPFKLIFSSTIQKRL.....LLTRDRFVLKVTAFRVNEYV 341

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79: \*  
1: PIR1: \*  
2: PIR2: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	946	52.7	226	E96531	hypothetical prote
2	229	12.8	190	T02015	hypothetical prote
3	227.5	12.7	461	T05979	hypothetical prote
4	182.5	10.2	308	T49920	hypothetical prote
5	172	9.6	286	H84537	hypothetical prote
6	172	9.6	389	T04786	hypothetical prote
7	161.5	9.0	352	T05106	hypothetical prote
8	145	8.1	344	T05104	hypothetical prote
9	141.5	7.9	210	D96531	hypothetical prote
10	141	7.9	283	T45871	hypothetical prote
11	140.5	7.8	461	T05101	hypothetical prote
12	137.5	7.7	490	T01545	hypothetical prote
13	134	7.5	512	T05107	hypothetical prote
14	133.5	7.4	984	G86393	protein T24P13.6 [
15	133.5	7.4	1440	C84639	hypothetical prote
16	133	7.4	899	H84639	hypothetical prote
17	132	7.3	493	T05105	hypothetical prote
18	118	6.6	478	T05102	hypothetical prote
19	116.5	6.5	851	F84639	hypothetical prote
20	111	6.2	608	AE2531	hypothetical prote
21	110.5	6.2	1021	T05108	hypothetical prote
22	107.5	6.0	2783	T41948	alpha-fetoprotein
23	104	5.8	497	T14433	reproductive meris
24	102.5	5.7	134	E86301	F19K19.7 protein -
25	102.5	5.7	232	T22698	hypothetical prote
26	101.5	5.7	243	TJ0788	tonB protein - Kle
27	101.5	5.7	482	T22981	hypothetical prote
28	101	5.6	1507	T18544	alpha-2-macroglobu
29	99	5.5	298	T37251	homeobox protein c

30	98	5.5	586	2	T29657	hypothetical prote
31	98	5.5	720	2	T26819	hypothetical prote
32	97.5	5.4	1082	2	T15269	hypothetical prote
33	97	5.4	252	2	A00267	tonB protein [impo
34	97	5.4	525	2	T05109	hypothetical prote
35	96.5	5.4	297	2	A84767	hypothetical prote
36	96	5.3	852	2	T06310	hypothetical prote
37	95.5	5.3	616	2	G96511	hypothetical prote
38	95	5.3	673	1	S73444	MG032 homolog B01
39	95	5.3	932	2	T45894	hypothetical prote
40	94.5	5.3	1003	2	T16740	hypothetical prote
41	94	5.2	1008	2	S38003	translation elonga
42	92.5	5.2	599	2	T00948	hypothetical prote
43	92	5.1	541	2	C81357	probable secreted
44	92	5.1	682	2	G84639	hypothetical prote
45	92	5.1	823	2	T09882	heat shock protein

ALIGNMENTS

RESULT 1

E96531

hypothetical protein F13F21.8 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: E96531

R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E96531

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-226 <STO>

A:Cross-references: UNIPROT:Q9XIB5; GB:AE005173; NID:G5430753; PIDN:AAD43153.1; GSPDB:G

C:Genetics:

A:Gene: F13F21.8

A:Map position: 1

Query Match 52.7%; Score 946; DB 2; Length 226;  
Best Local Similarity 81.2%; Pred. No. 3.3e-69;  
Matches 190; Conservative 14; Mismatches 18; Indels 12; Gaps 5;

QY 110 MDSAHNF-KEARLPEDLEDAEVIFFSSVYPSLPSTVPANKGY-ASSAIQILFTCP 167

Db 3 MDSAQNFNKARLPEDLPKDAKVIYPSN-----PESTEFVNKGYGGSTAIQSFFK-E 55

QY 168 VKAEPTPTPKPKKGRKKKNADPEEINSAPRDDDDPENRSKFYESASARKTSTAER 227

Db 56 SKAEE---TPVKLKKRKKKKKNPEEYNSSTPGDDSENRSKFYESASARKTSTAER 112

QY 228 ERANNAKTFFTPNPFFRVLRPSLYRGICIMYLPSPGFAEKYLSGISGFIKVQLAEKQWP 287

Db 113 ERAVNAKTFFTPNPFFRVLRPSLYRGICIMYLPSPGFAEKYLSGISGFIKLQGEKQWP 172

QY 288 VRCLYKAGRAKPSQGWYFTELENNIGEGDVCVFELLRTDRFVLKVTAFRVNEYV 341

Db 173 VRCLYKAGRAKPSQGWYFTELENNIGEGDVCVFELLRTDRFVLKVTAFRVNEYV 226

RESULT 2

T02015

hypothetical protein T15B16.18 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 09-Jul-2004



A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: H84537

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-286 <STO>

A:CROSS-references: UNIPROT:Q9XIG9; GB:AE002093; NID:94678216; PIDN:AAD26962.1; GSPDB:GN

C:Genetics:

A:Gene: At2g16210

A:Map position: 2

C:Superfamily: Arabidopsis thaliana hypothetical protein F17114.30

Query Match 9.6%; Score 172; DB 2; Length 286;

Best Local Similarity 21.3%; Pred. No. 2e-06;

Matches 72; Conservative 56; Mismatches 128; Indels 82; Gaps 13;

QY 6 FHLKIFSSSTI-DEKRLRVPDKFVSKFD-ELSVAVALTVPDGHVVRVGLRKADNKIWFOD 63

DB 19 FFAVQVSINVSSENKALPHDSRSTDELKSRKMKIRAQWNSWEVGLSK-NPRIFYME 77

QY 64 --GWQEFVDYRISIRIGYLLIFRYEGNSAFSVYIFNLHSEINYNHSTGLMDSAHNHFKRAR 121

DB 78 KSGWEKFRVDNALGNSSELLTFTHKGWGHFTVNIPLKDGKE-----NMQPPQSRSRFFAS 130

QY 122 LPEDLEDEDAEVIFPSSVPSLPSTVPANKYASSAIQTLFTGPVKAEEPTTPKPKP 181

DB 131 SKQEENDIKEEVV-----VSSNRGQITAA-----E 155

QY 182 KGRKKKNADPEINSSAPRDDDDPENRSKFYESASARKETVTAERERAINAKTF---- 237

DB 156 SKGRKD-----NLGK-----RAKESQSKTKETKVVRRARSVDAGAS 191

QY 238 BPTNPFPRVLRPSYLRGIMYLPSPGFAEKYLGSIGFIKQV--LAERQWVRCILYKAG 295

DB 192 SSTAATAFTILFQGYL--VFLRIPNSVSKQVPDEKTVFKIHHPNGKSNVVLIRG 248

QY 296 RAKFSQGWTEFTLENNLGGDVCVFELLTRDFVLKVT 333

DB 249 --AFSGWRVRVKEVPLAVGDTCKFTFKPKELLVVS 284

#### RESULT 6

T04786

hypothetical protein F10M10.170 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C:Accession: T04786

R:Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Hoheisel, J.; Mewes, H.W.; Mayer,

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15384

A:Accession: T04786

A:Molecule type: DNA

A:Residues: 1-389 <BEV>

A:CROSS-references: UNIPROT:Q9S205; EMBL:AL035521

A:Experimental source: cultivar Columbia; BAC clone F10M10

C:Genetics:

A:Map position: 4

A:Introns: 29/3; 154/1; 203/1; 276/1; 306/3

A:Note: F10M10.170

Query Match

Best Local Similarity 9.6%; Score 172; DB 2; Length 389;

Matches 86; Conservative 51; Mismatches 154; Indels 118; Gaps 16;

QY 1 MRPFPFKLIFSTIQEKLRLVPDKFVSKFDELSVAVALTVPDGHVVRVGLRKADNKI 60

DB 12 LPRPFF--TVFVSHFSSFMWIPVSVYDHIHRFPKTVILRGPGGCKWVATEIKDDVYL 69

QY 61 FQDGHQEFVDYRISIRIGYLLIFRYEGNSAFSVYIF-----N 96

DB 70 FQGHKPKFRVDNLNDGDLFTAYNGAHIFEVISIRFGYDACKIEVTELEBEREDSVIS 129

QY 97 LSHSEINYNHSTGLMDSAHNHFKRAR-----LPEDLEDEDAEVIFPSSVPSLPST 148

DB 130 LSSEDT---TGAKSEMKNVTPEGRDKGSKVVEVEDSDDBEE-----DSVYSESEBETE 182

QY 149 VPANKGYASSAIQTLFTGPVKAEEPTTPKPK--KGRKK----- 187

DB 183 TDTSEF-----KVAKPT-----IPKQKKGKKKQVQVSSDEDEBEDSDS 225

QY 188 -----KNADPEINSSAPRDDDDPENRS-----KFYESASARKTV--TABERER 229

DB 226 DYIETFGQLDIENGISEEDSSYAPDKEDTATASFVKPKVANKVANLKRKKVDPMIKNP 285

QY 230 AINAAKT--FEPT--NPFPRVLRPSYLRGIMYLPSPGFAEKYLGSIGSIFKVLQAEKQ 285

DB 286 YLDDPKNIHFETNVKNRLYELVHAQLVKDYCLRF-----GDYNYIDRFGLSATAK 339

QY 286 WPRV--CLYKAGRAKFSQGWTEFTLENNLGGDVCVFELLTRDFVLKV 332

DB 340 WKDQRCVTKR-----WPRICKRNKLKEDRILCELLRKTGTFYAI 379

#### RESULT 7

T05106

hypothetical protein F28M20.170 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C:Accession: T05106

R:Bevan, M.; Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Hoheisel, J.; Mewes,

submitted to the Protein Sequence Database, November 1998

A:Reference number: Z15398

A:Accession: T05106

A:Molecule type: DNA

A:Residues: 1-352 <BEV>

A:CROSS-references: UNIPROT:Q9SB80; EMBL:AL031004

A:Experimental source: cultivar Columbia; BAC clone F28M20

C:Genetics:

A:Map position: 4

A:Introns: 26/3; 125/1; 166/3

A:Note: F28M20.170

C:Superfamily: Arabidopsis thaliana hypothetical protein F28M20.180

Query Match

Best Local Similarity 9.0%; Score 161.5; DB 2; Length 352;

Matches 69; Conservative 49; Mismatches 102; Indels 123; Gaps 15;

QY 5 PFHKLI--FSSITQEKRLRLVPDKFVSKF---KDELSVAVALTVPDGHVVRVGLRKADNKI 59

DB 14 FQPTILPGFKS-----HIKIPVKFSTKIEGHEGNTVLRSPSRRTWKV--KMEGHKL 66

QY 60 WPDQGWQFVDYRISIRIGYLLIFRYEGNSAFSVYIFNLHSEINYNHSTGLMDSAHNHFKR 119

DB 67 --TEGWKEFVEAHLRLVGDVVFVFKHGDMLFHTVTAIGPSCCEVQY----- 109

QY 120 ARLFEDLEDEDAEVIFPSSVPSLPSTVPANKGYASSAIQTLFTGPVKAEEPTTPKI 179

DB 110 -----APSRSHD----- 116

QY 180 PKKGRKKKNADPEINSSAPRDDDDPENRSKFYESASARKTVTAERERAINAKTFEP 239

DB 117 -----RNESEDEIGSS-----RNEKILIE-----ENVKTEPDQFSP 147

QY 240 TNPFRVLRPSYLRGIMYLPSPGFAEKYLGSIGS-GFIKVLQAEKQ-----WPVRC-LYKA 294

DB 148 DLTCFSQSQTASNLTRD-LVGIPRDFAKRY--GLNIGRHEIVLMBDEGNTSESEVSKYS 204

QY 295 GRAKFSQGWTEFTLENNLGGDVCVFELLTRDFVLKVTAFRV 337

DB 205 GRVFIAGGWTSLCTANKLEVGDSTCFKLRLNP-----KIPVFL 243

#### RESULT 8

T05104

hypothetical protein F28M20.150 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004



submitted to the Protein Sequence Database, November 1998

A:Reference number: Z15398  
A:Accession: T05107  
A:Molecule type: DNA  
A:Residues: 1-461 <BEV>  
A:Cross-references: UNIPROT:O81778; EMBL:AL031004  
A:Experimental source: cultivar Columbia; BAC clone F28M20  
C:Genetics:  
A:Map position: 4  
A:Introns: 26/3; 125/1; 163/3; 245/3; 316/3  
A:Note: F28M20.120  
C:Superfamily: Arabidopsis thaliana hypothetical protein F28M20.180

Query Match 7.8%; Score 140.5; DB 2; Length 461;  
Best Local Similarity 21.5%; Pred. No. 0.0014;  
Matches 81; Conservative 50; Mismatches 151; Indels 95; Gaps 18;

QY 5 FFKLI--FSSITQKRLVDPKFSKFP---KDELSVALVTPDGHVVRV---GLRKAD 56  
DB 14 FFQPLLPGFQS-----NLKIPVNYFSEHTGHEGKTVTLRTDASERTWEVRMEGHR--- 65  
QY 57 NKIWFQDQWQFVDRYSIRIGVLLIFRYEGNSAFSVIENLHSEINYN-----HSTGLMD 111  
DB 66 ----LTEGKWEFVEADHLRIGDFVVRHGDVWFVHTALGPSCCELIQYQSSRHEEG-EE 120  
QY 112 SAHNHFKRARLPEDLEDEDAEV-----IFPSSVYPSLPPESTVPANKGYASSAIQTFTG 166  
DB 121 SCENEISEKEGENVQKESDKSSDLNCFPSQSVTHSNISRDVAVPRDFV----- 170  
QY 167 PVKAEPTTPKIPKGRGKKQKNADPEEINSSAPRDDDDPENRSKFYESAS---ARKRTVT 223  
DB 171 -----KRSFGSKGR---HEIVLMEEGKSWESSEVKYMSGAVYLVGWTTF 213  
QY 224 AEERERAINAAKTF---EPTWPFPRVVLVRLPSVLYRGCIWYLPSPGAEKYLGSIGFI-- 277  
DB 214 CTENKLDVGDSTFKLLQAKTPVQLCSRTK-----HPLSFTK-----VNGLINP 260  
QY 278 -KVQLAEK---QWPVRLCYKAGRAKFSQSG---WYEFTELENNLGEQDVCFELLR----- 324  
DB 261 GKILLVDKDRAEWSMMLKVDKRGAVYIIGNDWKSCFAANEVGAESLALELIQGGVLLN 320  
QY 325 ---TRDFVLKVTAFRNE 339  
DB 321 QITTCFQWQPSFKAED 337

## RESULT 12

T01545  
hypothetical protein A\_IG005110.15 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 09-Jul-2004  
A:Reference number: Z14347  
A:Accession: T01545  
R:Andrews, S.  
submitted to the EMBL Data Library, July 1997  
A:Description: The sequence of A. thaliana IG005110.  
A:Reference number: Z14347  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-490 <AND>  
A:Cross-references: UNIPROT:O23076; EMBL:AF013293; NID:g2252823; PID:g2252835  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Map position: 4  
A:Introns: 112/3; 216/1; 256/3; 391/3  
A:Note: A\_IG005110.15  
C:Superfamily: Arabidopsis thaliana hypothetical protein F28M20.180

Query Match 7.7%; Score 137.5; DB 2; Length 490;  
Best Local Similarity 23.4%; Pred. No. 0.0026;  
Matches 81; Conservative 42; Mismatches 158; Indels 65; Gaps 16;

QY 22 VPDKFVSKF---KDELSVALVTPDGHVVRVGLRKADNKIWFQDQWQFVDRYSIRIGY 78

Db 6 IPVAFFLYIEGRYEQKTAKLASDASKRTWEY---KIDGQ-RLTDGKWEFAVSHDLRIGD 61  
QY 79 LLIFRYEGNSAFSVIENLHSEINVHSTGLM-----DSAHNHFKRARLFE 124  
Db 62 IVVFRQESDLAFHVTLGLPSCCGIQYGSCSVKNNLGDGELCSVSGSSRNVRFPGRFVRD 121  
QY 125 DLEDEDAEVI-----FPSSVYPSLPPESTVPANKGYASSAIQTFTGVPKABEPT 174  
Db 122 NGVVGSGEIVLMNEKGRSWNLRQKPS---NGTVVRGWWGWSFC---DANGLKAGD-N 173  
QY 175 PTPKIPKRG-----RKKNADPEEINS---SAPRDDDDPENRSKFYESASAKRTVTABE 226  
Db 174 YTFKLIRAGTLVLRLLPNEPKEANEVSLPEPESDAERNLEKIQKKEV-KKNVTRE- 231  
QY 227 RERAINAAKTRPTNPFPRVVLVRLPSVLYRGCIWYLPSPGAEKYLGSIGI-----SGFIKVQLA 282  
Db 232 -----AESQDSQSCFVANVSPSSL-RYDTLYLPKRFMRB--NGVDKRCGEMILINEK 281  
QY 283 EKQWVRCLYK--AGRAKFSQGYEFTLENNLGEQDVCFELLRT 326  
Db 282 GKSWTLDLKVKSSTSLIKGRWSFCSANGLRAGSIITLKLKKR 327

RESULT 13  
T05107  
hypothetical protein F28M20.180 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
A:Accession: T05107  
R:Bevan, M.; Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Hoheisel, J.; Mewes, J.  
submitted to the Protein Sequence Database, November 1998  
A:Reference number: Z15398  
A:Accession: T05107  
A:Molecule type: DNA  
A:Residues: 1-512 <BEV>  
A:Cross-references: UNIPROT:Q9SB79; EMBL:AL031004  
A:Experimental source: cultivar Columbia; BAC clone F28M20  
C:Genetics:  
A:Map position: 4  
A:Introns: 26/3; 184/3; 322/3; 402/3  
A:Note: F28M20.180  
C:Superfamily: Arabidopsis thaliana hypothetical protein F28M20.180

Query Match 7.5%; Score 134; DB 2; Length 512;  
Best Local Similarity 19.9%; Pred. No. 0.0053;  
Matches 69; Conservative 47; Mismatches 111; Indels 120; Gaps 12;

QY 1 MPRP----FFHKLIFSSTIQKRLVDPKFSV-SKFKDEL-SVAVALTVDPDGHVVRVGLRK 54  
Db 6 IPSPTNKAF---IIDLSGQKSNPIPTPTFIWNHFNKIQSTNMKLT-----SD 51  
QY 55 ADNKIW-----PDQWQFVDRYSIRIGVLLIFRYEGNSAFSVIENLHSEINYNYS 106  
Db 52 ASDRNWDVYKLDGARFAGWKDFSVSHSVRDDDLSFRHDGGMVHVSPFGRSFSQILIS 111  
QY 107 TGLMDSAHNHFKRARLPEDLEDEDAEVIFPSSVYPSLPPESTVPANKGYASSAIQTFTG 166  
Db 112 SSTSDDDDDD---ERTVDDDEDDDDVGGDDDDNSI-----SEDDFCCKKISS----- 153  
QY 167 PVKABEPTTPKIPKGRGKKKNADPEE-----INSSAPRDDDDPENRSKFYESASAKR 220  
Db 154 -----KRRARKETESSSDKSYLVHAHTVTPSSLLRDNMCVLSKPARSNGLDR 199  
QY 221 TVTAERERAINAAKTRPTNPFPRVVLVRLPSVLYRGCIWYLPSPGAEKYLGSIGFIKVQ 280  
Db 200 BCEIDLDRD----- 208  
QY 281 LAEKOMP--VRCLYKAGRAKFSQGYEFTLENNLGEQDVCFELLRT 325  
Db 209 -HEKSWTLLLRHNKKTGQAFMRGGRWSFCRNNGIKAGSICRPFKLVS 254

Mon Jan 3 13:13:31 2005

RESULT 14  
G86393  
protein T24P13.6 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: G86393  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.; Huizar, L.  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: G86393  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-984 <STO>  
A:Cross-references: UNIPROT:Q9LQV4; GB:AE005172; NID:g9295721; PIDN:AAF87027.1; GSPDB:GN  
C:Genetics:  
A:Gene: T24P13.6  
A:Map position: 1

Query Match 7.4%; Score 133.5; DB 2; Length 984;  
Best Local Similarity 21.4%; Pred. No. 0.014;  
Matches 81; Conservative 46; Mismatches 151; Indels 101; Gaps 18;  
QY 5 FFHKLIFSSTIOEKRLVPDKFVSKF---KDELSVAVALTPDGHVWRVGLRKADNKIW- 60  
DB 15 FPQFIL---TESRTHLNIIPVAFPSKHVEGRNQNTVILR-----SDASDKTWL 60  
QY 61 -----FDGQWQFVDYRISIRIGYLLIPRYEGNSAFSYIIFNLSHSYNHSTGLMDSA 113  
DB 61 VKMDGLKLTGDWEDFAFADHLATGDIVFRLGEMVHFVTAALGPSCCEIQYHT-----SS 115  
QY 114 HNHFKARLPEDLEDAEVIFFSSVYPSPLPSTVPAN--KGYASSAIQTLFTGPVKA- 170  
DB 116 HN-----INDDRN-----DQINIASRNSRKYKPKRKSVESSLDHSRFAKVSAM 161  
QY 171 ---EEPTPTP-KIPKRGKKNADPEEINSSAPRD-----DDPENRSKFYESA----- 215  
DB 162 CLSNDRLYPLSFARLNGLNKINSKYLQNEGRSKVLRLRDKSGWQTFVQSGWRRFC 221  
QY 216 -----SA-----RKTVTABERERAINAAKTPEPTNPFRRVVLPSYLY 254  
DB 222 SENGIRQGVTFKLVKRSAPPVIRLCRAKAPKQRSV---AEYSSDHSCFEGSVTPSSL- 277  
QY 255 RGCIMVLPSCGFA-----EKVLSGISGFIKVLQAEKO---WPVRCLYKAGRAKFSQGWYEF 306  
DB 278 RNDLLYLPKRSFVNSNRLDKRCS-----EIVLKNQGVKRWPLVLKRFKSVTYLPKGWTSF 331  
QY 307 TLENLNGEGDVCVFELLRT 325  
DB 332 CQVNRKAGDSFKFLVGT 350

RESULT 15  
C84639  
hypothetical protein At2g24650 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: C84639  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: C84639  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-1440 <STO>  
A:Cross-references: UNIPROT:Q9SJA0; GB:AE002093; NID:g4572680; PIDN:AAD23895.1; GSPDB:GN  
C:Genetics:  
A:Gene: At2g24650  
A:Map position: 2

Query Match 7.4%; Score 133.5; DB 2; Length 1440;  
Best Local Similarity 22.4%; Pred. No. 0.024;  
Matches 87; Conservative 52; Mismatches 127; Indels 123; Gaps 20;  
QY 5 FFHKLIFSSTIOEKRLVPDKFVSKFDELSVAVALTPDGHVWRVGLRKADNKIWFQDG 64  
DB 10 FFHTLVPSF---HTHLMIPEDFPSEYIEGRSVAELKLDKFSKSWVKL--SDRRI--TDG 62  
QY 65 WQFVDYRISIRIGYLLIPRYEGNSAFSYIIFNLSHSYNHSTGLMDSAHN-----HFKRA 120  
DB 63 WEEFVWANDFRIGDVVAFRYVGN-----LVFHVSNLGNPNY-----EIEHNDLPKEKKA 112  
QY 121 RLFEDLEDAEVIFFSSVYPSPLPSTVPANKGYASSAIQTLFTGPVKA-EEPTPTPKI 179  
DB 113 K-----TNSEADAVSS-----SSADKSCFMAITALTDTLTYL 149  
QY 180 P-----KRGKKNAD-----PEEINSSAPRDDDPENRSKFYESASAR----- 218  
DB 150 PLHFTSANGLTRKNREIILTDGERSRVLDLRFD--ESSGTFYISRGWRNFCDENGQKAG 207  
QY 219 -----KRTVTAE-----ERERAI----- 231  
DB 208 GFFLKLVLGKETLVLVSFCPTESINGEENTRDSKDECSLDSLMNIVEKKYIPKPRG 267  
QY 232 NAAKTPEPTN-PFRRVVLPSYLYRGCIMYLPSCGFAEKYLSGISGFIKVLQAEK---OWP 287  
DB 268 SPYSSYSPSHKQFVTFPLPDYARIGKLS-LSAPFVRE--NGINKPGEICLLDKHGRKWL 324  
QY 288 VRCLYKA-GRAKFSQGWYEFLENLGE 315  
DB 325 TSLLLDSKGTMSLKGKWKFEVKANSLETG 353

Search completed: December 30, 2004, 06:11:10  
Job time : 25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2004, 12:08:07 ; Search time 79 Seconds  
(without alignments)

1548.439 Million cell updates/sec

Title: US-10-088-187A-11

Perfect score: 1796

Sequence: 1 MRPFPFKLIFSTIQKRL.....LLRTDFVLKVTAFRNEYV 341

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1796	100.0	341	4 AAB35491	Aab35491 A thalian
2	1796	100.0	341	8 ADN73923	Adn73923 Thale cre
3	1789	99.6	341	3 AAG22414	Aag22414 Arabidops
4	1711.5	95.3	448	3 AAG54185	Aag54185 Arabidops
5	1645	91.6	377	3 AAG54186	Aag54186 Arabidops
6	1217	67.8	232	3 AAG22415	Aag22415 Arabidops
7	1217	67.8	232	3 AAG54196	Aag54196 Arabidops
8	1170	65.1	284	3 AAG54187	Aag54187 Arabidops
9	953	53.1	182	3 AAG17883	Aag17883 Arabidops
10	946	52.7	224	3 AAG58317	Aag58317 Arabidops
11	946	52.7	224	3 AAG61371	Aag61371 Arabidops
12	946	52.7	226	3 AAG61370	Aag61370 Arabidops
13	946	52.7	226	3 AAG58316	Aag58316 Arabidops
14	940	52.3	224	3 AAG17136	Aag17136 Arabidops
15	940	52.3	226	3 AAG17135	Aag17135 Arabidops
16	293.5	16.3	346	7 ADC03416	Adc03416 Rice flow
17	226.5	12.6	200	3 AAG10757	Aag10757 Arabidops
18	226.5	12.6	206	3 AAG10756	Aag10756 Arabidops
19	226.5	12.6	214	3 AAG10755	Aag10755 Arabidops
20	221	12.3	243	3 AAG54198	Aag54198 Arabidops
21	221	12.3	246	3 AAG54197	Aag54197 Arabidops
22	215.5	12.0	209	3 AAG54179	Aag54179 Arabidops
23	199	11.1	291	3 AAG47303	Aag47303 Arabidops
24	199	11.1	300	3 AAG47302	Aag47302 Arabidops
25	196	10.9	205	3 AAG54199	Aag54199 Arabidops

26	195.5	10.9	292	3	AAG06195	Aag06195 Arabidops
27	195.5	10.9	301	3	AAG06194	Aag06194 Arabidops
28	190.5	10.6	171	3	AAG54180	Aag54180 Arabidops
29	182.5	10.2	167	3	AAG54181	Aag54181 Arabidops
30	180	10.0	270	3	AAG06196	Aag06196 Arabidops
31	172	9.6	469	3	AAG47279	Aag47279 Arabidops
32	172	9.6	491	3	AAG47278	Aag47278 Arabidops
33	172	9.6	500	3	AAG47277	Aag47277 Arabidops
34	160.5	8.9	250	3	AAG47304	Aag47304 Arabidops
35	145.5	8.1	230	7	ADC03364	Adc03364 Rice flow
36	139.5	7.8	286	3	AAG06859	Aag06859 Arabidops
37	139.5	7.8	286	3	AAG54115	Aag54115 Arabidops
38	139.5	7.8	302	3	AAG06858	Aag06858 Arabidops
39	139.5	7.8	302	3	AAG54114	Aag54114 Arabidops
40	129	7.2	262	3	AAG06860	Aag06860 Arabidops
41	129	7.2	262	3	AAG54116	Aag54116 Arabidops
42	129	7.2	280	3	AAG60813	Aag60813 Arabidops
43	129	7.2	280	3	AAG06242	Aag06242 Arabidops
44	129	7.2	280	3	AAG60886	Aag60886 Arabidops
45	129	7.2	280	3	AAG60888	Aag60888 Arabidops

ALIGNMENTS

RESULT 1  
AAB35491  
ID AAB35491 standard; protein; 341 AA.  
XX AC AAB35491;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE A thaliana VRN1.  
XX  
KW VRN1; vernalisation; flowering; crop.  
XX  
OS Arabidopsis thaliana.  
XX  
FH Key Location/Qualifiers  
FT Domain 2..94  
FT Region /label= B3\_DNA\_binding\_domain  
FT Region 95..238  
FT Domain /label= region\_2  
FT Domain 239..332  
FT Domain /label= B3\_DNA\_binding\_domain  
XX  
PN WO200121822-A1.  
XX  
PD 29-MAR-2001.  
XX  
PP 13-SEP-2000; 2000WO-GB003525.  
XX  
PR 17-SEP-1999; 99GB-00022071.  
XX  
PA (PLAN-) PLANT BIOSCIENCE LTD.  
XX  
PI Dean C, Levy YY;  
XX  
DR WPI; 2001-273467/28.  
XX  
DR N-PSDB; AAF62446.  
XX  
PT Novel VRN1 polynucleotide sequence encoding a polypeptide which alters  
PT vernalization response of plant in which VRN1 nucleic acid is expressed,  
XX useful for influencing and assessing vernalization phenotype of plants.  
XX  
PS Claim 25; Fig 7; 91pp; English.  
XX  
CC The present invention provides the protein and coding sequences of  
CC Arabidopsis thaliana VRN1. This protein is capable of altering the  
CC vernalisation responses of a plant. Also provided are a number of PCR  
CC primers used to isolate the sequences. The sequences are useful in the  
CC production of crop plants, where they are able to control the timing of

CC flowering, the duration of vernalisation required, the optimum  
CC temperature, or even eliminate the need for vernalisation completely. The  
CC present sequence is the VRN1 protein

XX SQ Sequence 341 AA;

Query Match 100.0%; Score 1796; DB 4; Length 341;  
Best Local Similarity 100.0%; Pred. No. 1.4e-167;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MPRPFHKLIFSSTIOEKRLRVPDKFVSKFDLSVAVALTPDGHVVRVGLRKADNKIW 60  
DB 1 MPRPFHKLIFSSTIOEKRLRVPDKFVSKFDLSVAVALTPDGHVVRVGLRKADNKIW 60  
QY 61 FQDQWQEFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLHSHSEINHYSTGLMDSAHNHFKRA 120  
DB 61 FQDQWQEFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLHSHSEINHYSTGLMDSAHNHFKRA 120  
QY 121 RLFEDLEDEDAEVIFFPSVYPSPLPESTVPANKGYASSAIOTLFTGPVKAEPTPTPKIP 180  
DB 121 RLFEDLEDEDAEVIFFPSVYPSPLPESTVPANKGYASSAIOTLFTGPVKAEPTPTPKIP 180  
QY 181 KRGKKNADPEEINSSAPRDDDPENRSKFYESASAKRTVTABERERAINAAKTPEPT 240  
DB 181 KRGKKNADPEEINSSAPRDDDPENRSKFYESASAKRTVTABERERAINAAKTPEPT 240  
QY 241 NPFFRVLRPSLYRGCIWYLPSPGFAEKYLSGISGFIKVQLAEKQWVRCCLYKAGRAKFS 300  
DB 241 NPFFRVLRPSLYRGCIWYLPSPGFAEKYLSGISGFIKVQLAEKQWVRCCLYKAGRAKFS 300  
QY 301 QGWYEFTLENNLGBGDVCFELLRTDFVLKVTAFRVNEYV 341  
DB 301 QGWYEFTLENNLGBGDVCFELLRTDFVLKVTAFRVNEYV 341

RESULT 2

ADN73923  
ID ADN73923 standard; protein; 341 AA.

XX AC ADN73923;

XX DT 15-JUL-2004 (first entry)

XX DE Thale cress protein repressed in E2Fa/Dpa expressing plants Seqid 1818.

XX KW plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;

XX KW animal feed product; thale cress; cell wall biosynthesis;

XX KW nitrogen metabolism; carbon metabolism.

XX OS Arabidopsis thaliana.

XX PN WO2004035798-A2.

XX PD 29-APR-2004.

XX PF 20-OCT-2003; 2003WO-EP011658.

XX PR 18-OCT-2002; 2002EP-00079408.

XX PA (CROP-) CROPDESIGN NV.

XX PI Inze D, De Veylder L, Vlieghe K;

XX DR WPI: 2004-348466/32.

XX DR N-PSDB; ADN73922.

XX Altering plant characteristics, useful for producing plants for enzyme or  
PT pharmaceutical production comprises modifying in a plant, expression of  
PT one or more nucleic acids and/or modifying level or activity of one or  
PT more proteins.

XX Claim 1; SEQ ID NO 1818; 134pp; English.

CC This invention relates to a novel method for altering one or more plant  
CC characteristics. Specifically, it refers to identifying genes that are up  
CC - or down-regulated in transgenic plants overexpressing the heterodimeric  
CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to  
CC alter plant characteristics accordingly. The present invention describes  
CC generating transgenic plants for the production of growth regulators,  
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where  
CC the altered plant characteristics are selected from increased yield or  
CC biomass, enhanced survival capacity, stress tolerance, plant architecture  
CC or physiology, altered endoreplication, biochemistry, signal  
CC transduction, storage lipid mobilisation and/or altered photosynthesis,  
CC each relative to the corresponding wild type plants. Accordingly, these  
CC sequences can also be useful as positive or negative selectable markers  
CC during transformation of cells or tissues. The identified genes play a  
CC role in a variety of biological processes such as DNA replication, cell  
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as  
CC transcription factors. This polypeptide sequence is thale cress protein  
CC expressed by a gene repressed 1.3 fold or more in plants overexpressing  
CC the E2Fa/Dpa transcription factor, given in an exemplification of the  
CC invention.

XX SQ Sequence 341 AA;

Query Match 100.0%; Score 1796; DB 8; Length 341;  
Best Local Similarity 100.0%; Pred. No. 1.4e-167;  
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPRPFHKLIFSSTIOEKRLRVPDKFVSKFDLSVAVALTPDGHVVRVGLRKADNKIW 60

DB 1 MPRPFHKLIFSSTIOEKRLRVPDKFVSKFDLSVAVALTPDGHVVRVGLRKADNKIW 60

QY 61 FQDQWQEFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLHSHSEINHYSTGLMDSAHNHFKRA 120

DB 61 FQDQWQEFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLHSHSEINHYSTGLMDSAHNHFKRA 120

QY 121 RLFEDLEDEDAEVIFFPSVYPSPLPESTVPANKGYASSAIOTLFTGPVKAEPTPTPKIP 180

DB 121 RLFEDLEDEDAEVIFFPSVYPSPLPESTVPANKGYASSAIOTLFTGPVKAEPTPTPKIP 180

QY 181 KRGKKNADPEEINSSAPRDDDPENRSKFYESASAKRTVTABERERAINAAKTPEPT 240

DB 181 KRGKKNADPEEINSSAPRDDDPENRSKFYESASAKRTVTABERERAINAAKTPEPT 240

QY 241 NPFFRVLRPSLYRGCIWYLPSPGFAEKYLSGISGFIKVQLAEKQWVRCCLYKAGRAKFS 300

DB 241 NPFFRVLRPSLYRGCIWYLPSPGFAEKYLSGISGFIKVQLAEKQWVRCCLYKAGRAKFS 300

QY 301 QGWYEFTLENNLGBGDVCFELLRTDFVLKVTAFRVNEYV 341

DB 301 QGWYEFTLENNLGBGDVCFELLRTDFVLKVTAFRVNEYV 341

RESULT 3

AAG22414  
ID AAG22414 standard; protein; 341 AA.

XX AC AAG22414;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 25333.

XX KW protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149923P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 27-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.

29-SEP-1999;	99US-0156596P.	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 69062.
04-OCT-1999;	99US-01571117P.	XX	Protein identification; signal transduction pathway; metabolic pathway;
05-OCT-1999;	99US-0157753P.	KW	hybridisation assay; genetic mapping; gene expression control; promoter;
06-OCT-1999;	99US-0157865P.	KW	termination sequence.
07-OCT-1999;	99US-0158029P.	XX	Arabidopsis thaliana.
08-OCT-1999;	99US-0158232P.	OS	
12-OCT-1999;	99US-0158369P.	XX	
13-OCT-1999;	99US-0159293P.	PN	EPI033405-A2.
13-OCT-1999;	99US-0159294P.	XX	
13-OCT-1999;	99US-0159295P.	XX	
14-OCT-1999;	99US-0159329P.	XX	
14-OCT-1999;	99US-0159330P.	XX	
14-OCT-1999;	99US-0159331P.	XX	
14-OCT-1999;	99US-0159637P.	XX	
14-OCT-1999;	99US-0159638P.	XX	
18-OCT-1999;	99US-0159584P.	XX	
21-OCT-1999;	99US-0160741P.	XX	
21-OCT-1999;	99US-0160767P.	XX	
21-OCT-1999;	99US-0160768P.	XX	
21-OCT-1999;	99US-0160770P.	XX	
21-OCT-1999;	99US-0160814P.	XX	
21-OCT-1999;	99US-0160815P.	XX	
22-OCT-1999;	99US-0160980P.	XX	
22-OCT-1999;	99US-0160981P.	XX	
22-OCT-1999;	99US-0160989P.	XX	
25-OCT-1999;	99US-0161404P.	XX	
25-OCT-1999;	99US-0161405P.	XX	
25-OCT-1999;	99US-0161406P.	XX	
26-OCT-1999;	99US-0161359P.	XX	
26-OCT-1999;	99US-0161360P.	XX	
26-OCT-1999;	99US-0161361P.	XX	
28-OCT-1999;	99US-0161920P.	XX	
28-OCT-1999;	99US-0161992P.	XX	
28-OCT-1999;	99US-0161993P.	XX	
29-OCT-1999;	99US-0162142P.	XX	
Query Match 99.6%; Score 1789; DB 3; Length 341;			
Best Local Similarity 99.7%; Pred. No. 6.6e-167;			
Matches 340; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	MPRPFFHKLIFSTIQEKLRLVDPKFSVKDELSSVAVALTVPDGHVVRVGLRKADNKIW 60	
Db	1	MPRPFFHKLIFSTIQEKLRLVDPKFSVKDELSSVAVALTVPDGHVVRVGLRKADNKIW 60	
QY	61	FQDQGEFVDRYSIRGYLLIFRYEGNSAFSVYIFNLHSEINHYSTGLMDSAHNHFKRA 120	
Db	61	FQDQGEFVDRYSIRGYLLIFRYEGNSAFSVYIFNLHSEINHYSTGLMDSAHNHFKRA 120	
QY	121	RLFEDLEDEDAEVIFFSSVYPSPLPESTVPANKGYASSAIQTLFTGPVKAEEPTTPKIP 180	
Db	121	RLFEDLEDEDAEVIFFSSVYPSPLPESTVPANKGYASSAIQTLFTGPVKAEEPTTPKIP 180	
QY	181	KKGGRKKQADPEINSSAPRODDPENRSKFYESASARKRTVTAERERAINAAKTFFPT 240	
Db	181	KKGGRKKQADPEINSSAPRODDPENRSKFYESASARKRTVTAERERAINAAKTFFPT 240	
QY	241	NPFPRVVLPSYLRGCIWYLPSPGFAEKYLSGIFIKVQLAEKQWPVCLYKAGRAKPS 300	
Db	241	NPFPRVVLPSYLRGCIWYLPSPGFAEKYLSGIFIKVQLAEKQWPVCLYKAGRAKPS 300	
QY	301	QGWYEFTLENNLGEQDVCVFELLTRDFVLKVTAFRVNEYV 341	
Db	301	QGWYEFTLENNLGEQDVCVFELLTRDFVLKVTAFRVNEYV 341	
RESULT 4			
AAG54185			
ID	AAG54185 standard; protein; 448 AA.		
XX			
AC	AAG54185;		
XX			
DT	18-OCT-2000 (first entry)		
XX			

```
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140891P.
PR 30-JUN-1999; 99US-0141187P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143352P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149729P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.

PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 95.3%; Score 1711.5; DB 3; Length 448;
Best Local Similarity 83.8%; Pred. No. 4e-159;
Matches 332; Conservative 2; Mismatches 7; Indels 55; Gaps 1;

QY 1 MRPFFHKLIFSSTIQEKKL----- 20
Db 1 MRPFFHKLIFSSTIQEKKL----- 60
QY 21 -----RVDPKFKVSKDELSVAVALTPDGHVVRVGLRKADNKIWFQDGM 65
Db 61 IICFFISHTGMLVORVPDKFKVSKDELSVAVALTPDGHVVRVGLRKADNKIWFQDGM 120
QY 66 QBFVDYRSIRIGYLLIFRYEGNSAPSVYIFNLHSEINVHSTGLMDSAHNHFKRARLPED 125
Db 121 QEFVDYRSIRIGYLLIFRYEGNSAPSVYIFNLHSEINVHSTGLMDSAHNHFKRARLPED 180
QY 126 LEDEDAEVIFFSSVYPSLPSTVPANKGYASSAIQTTLFTGPVKAEEPTPTPKIPKQGR 185
Db 181 LEDEDAEVIFFSSVYPSLPSTVPANKGYASSAIQTTLFTGPVKAEEPTPTPKIPKQGR 240
QY 186 KKNADPEEINSSAPRDDDPPNRSKFYESASARKRTVTAERERAINAAKTPEPTNPFFR 245
Db 241 KKNADPEEINSSAPRDDDPPNRSKFYESASARKRTVTAERERAINAAKTPEPTNPFFR 300
```

Mon Jan 3 13:13:29 2005

```
QY 246 VVLRPSYLRGCMYLPSPGFAEKYLSGSGFIKVQLAEKQMPVRCLYKAGRAKFSQGWYE 305
Db 301 VVLRPSYLRGCMYLPSPGFAEKYLSGSGFIKVQLAEKQMPVRCLYKAGRAKFSQGWYE 360
QY 306 FTLENNLGEQDVCFVPELLTRDFVLKVTAFRVNEV 341
Db 361 FTLENNLGEQDVCFVPELLTRDFVLKVTYRGEANDVI 396

RESULT 5
AAG54186
ID AAG54186 standard; protein; 377 AA.
XX
AC AAG54186;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 69063.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 29-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132240P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 22-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140981P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
```

PR	11-AUG-1999;	99US-0148319P.	QY	81	IFRYEGNSAFSVYIFNLHSEINHYHSTGLMDSAHNHFKARLFEDLEDAEVIPSSVY	140
PR	12-AUG-1999;	99US-0148341P.	Db	65	IFRYEGNSAFSVYIFNLHSEINHYHSTGLMDSAHNHFKARLFEDLEDAEVIPSSVY	124
PR	13-AUG-1999;	99US-0148565P.	QY	141	PSPLPESTVPANKGYASSAIQTLFTGPVKAEBPTTPPKIPKGRGRKKQVADPEEINSSAP	200
PR	16-AUG-1999;	99US-0149368P.	Db	125	PSPLPESTVPANKGYASSAIQTLFTGPVKAEBPTTPPKIPKGRGRKKQVADPEEINSSAP	184
PR	17-AUG-1999;	99US-0149175P.	QY	201	RDDDPENRSKFYESAKRKTVTAEERERAINAAKTFEPTNPFRRVLRPSLYRCIMY	260
PR	18-AUG-1999;	99US-0149426P.	Db	185	RDDDPENRSKFYESAKRKTVTAEERERAINAAKTFEPTNPFRRVLRPSLYRCIMY	244
PR	20-AUG-1999;	99US-0149723P.	QY	261	LPSGFAEKYLSGSGFIKVQLAEKQMPVRCLYKAGRAKFSQGWYEFTLENNLGEQDVCVF	320
PR	20-AUG-1999;	99US-0149929P.	Db	245	LPSGFAEKYLSGSGFIKVQLAEKQMPVRCLYKAGRAKFSQGWYEFTLENNLGEQDVCVF	304
PR	23-AUG-1999;	99US-0149302P.	QY	321	ELLTRDFVLKVTAFRVNEYV	341
PR	23-AUG-1999;	99US-0149330P.	Db	305	ELLTRDFVLKTYRGEAVDYI	325
PR	25-AUG-1999;	99US-0150566P.				
PR	26-AUG-1999;	99US-0150884P.				
PR	27-AUG-1999;	99US-0151065P.				
PR	27-AUG-1999;	99US-0151066P.				
PR	27-AUG-1999;	99US-0151080P.				
PR	30-AUG-1999;	99US-0151303P.				
PR	31-AUG-1999;	99US-0151438P.				
PR	01-SEP-1999;	99US-0151930P.				
PR	07-SEP-1999;	99US-0152363P.				
PR	10-SEP-1999;	99US-0153070P.				
PR	13-SEP-1999;	99US-0153758P.				
PR	15-SEP-1999;	99US-0154018P.				
PR	16-SEP-1999;	99US-0154039P.				
PR	20-SEP-1999;	99US-0154779P.				
PR	22-SEP-1999;	99US-0155139P.				
PR	23-SEP-1999;	99US-0155486P.				
PR	24-SEP-1999;	99US-0155659P.				
PR	28-SEP-1999;	99US-0156458P.				
PR	29-SEP-1999;	99US-0156596P.				
PR	04-OCT-1999;	99US-0157117P.				
PR	05-OCT-1999;	99US-0157753P.				
PR	06-OCT-1999;	99US-0157865P.				
PR	07-OCT-1999;	99US-0158029P.				
PR	08-OCT-1999;	99US-0158232P.				
PR	12-OCT-1999;	99US-0158369P.				
PR	13-OCT-1999;	99US-0159293P.				
PR	13-OCT-1999;	99US-0159294P.				
PR	13-OCT-1999;	99US-0159295P.				
PR	14-OCT-1999;	99US-0159329P.				
PR	14-OCT-1999;	99US-0159330P.				
PR	14-OCT-1999;	99US-0159331P.				
PR	14-OCT-1999;	99US-0159637P.				
PR	14-OCT-1999;	99US-0159638P.				
PR	18-OCT-1999;	99US-0159584P.				
PR	21-OCT-1999;	99US-0160741P.				
PR	21-OCT-1999;	99US-0160767P.				
PR	21-OCT-1999;	99US-0160768P.				
PR	21-OCT-1999;	99US-0160770P.				
PR	21-OCT-1999;	99US-0160814P.				
PR	21-OCT-1999;	99US-0160815P.				
PR	22-OCT-1999;	99US-0160980P.				
PR	22-OCT-1999;	99US-0160981P.				
PR	22-OCT-1999;	99US-0160989P.				
PR	25-OCT-1999;	99US-0161404P.				
PR	25-OCT-1999;	99US-0161405P.				
PR	25-OCT-1999;	99US-0161406P.				
PR	26-OCT-1999;	99US-0161359P.				
PR	26-OCT-1999;	99US-0161360P.				
PR	26-OCT-1999;	99US-0161361P.				
PR	28-OCT-1999;	99US-0161920P.				
PR	28-OCT-1999;	99US-0161992P.				
PR	28-OCT-1999;	99US-0161993P.				
PR	29-OCT-1999;	99US-0162142P.				

Query Match 91.6%; Score 1645; DB 3; Length 377;  
Best Local Similarity 97.2%; Pred. No. 1.1e-152;  
Matches 312; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY	21	RVPDKFVSKFDLSVAVALTPDGHVWRVGLRKADNKIWFQDGHQWQEFVDYRSIRIGYLL	80
Db	5	RVPDKFVSKFDLSVAVALTPDGHVWRVGLRKADNKIWFQDGHQWQEFVDYRSIRIGYLL	64

RESULT 6  
AAG22415  
ID AAG22415 standard; protein; 232 AA.  
XX  
AC AAG22415;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 25334.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination assay.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 06-AUG-1999; 99US-0147432P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 09-AUG-1999; 99US-0148171P.  
PR 10-AUG-1999; 99US-0148319P.  
PR 11-AUG-1999; 99US-0148341P.  
PR 12-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 13-AUG-1999; 99US-0148688P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.

PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.  
  
Query Match 67.8%; Score 1217; DB 3; Length 232;  
Best Local Similarity 100.0%; Pred. No. 6.4e-111;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 110 MDSAHNHFKRARLFEDEDEDAEVIIPSSVYPSLPSTVPANKGYASSAIOTLFTGPVK 169  
DB 1 MDSAHNHFKRARLFEDEDEDAEVIIPSSVYPSLPSTVPANKGYASSAIOTLFTGPVK 60  
  
QY 170 ABEPTPTPKPKKGRKKKNADPEEINSSAPRDDDDPENRSKFYESASAKRTVTAEERER 229  
DB 61 ABEPTPTPKPKKGRKKKNADPEEINSSAPRDDDDPENRSKFYESASAKRTVTAEERER 120  
  
QY 230 AINAATFPTNPFPRVLRPSLYRGCIIMYLPSPGAEKYLSGISGFIKVQLAEKQMPVR 289  
DB 121 AINAATFPTNPFPRVLRPSLYRGCIIMYLPSPGAEKYLSGISGFIKVQLAEKQMPVR 180  
  
QY 290 CLYKAGRAKFSQGWYEFTLENNLGEQDVCVFELLRTDFVLKVTAPRVNEYV 341  
DB 181 CLYKAGRAKFSQGWYEFTLENNLGEQDVCVFELLRTDFVLKVTAPRVNEYV 232  
  
RESULT 7  
AAG54196  
ID AAG54196 standard; protein; 232 AA.  
AC AAG54196;  
XX  
XX  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 69077.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000BP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132485P.  
PR 05-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140821P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.







```
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145318P.
PR 27-JUL-1999; 99US-0145319P.
PR 28-JUL-1999; 99US-0145351P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 18-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.

Query Match 53.1%; Score 953; DB 3; Length 182;
Best Local Similarity 100.0%; Pred. No. 4e-85; 0; Indels 0; Gaps 0;
Matches 181; Conservative 0; Mismatches 0;

QY 161 QTLTGPVKAERTPTPKIPKRGKKKNADPEEINSSAPRDDDDPENRSKPFYESASARKR 220
    |||||
Db 2 QTLTGPVKAERTPTPKIPKRGKKKNADPEEINSSAPRDDDDPENRSKPFYESASARKR 61

QY 221 TVTAEERERAINAAKTPEPTNPFVRLRPSLYRGCMYLPSPGFAEKYLSGISGFIKVQ 280
    |||||
Db 62 TVTAEERERAINAAKTPEPTNPFVRLRPSLYRGCMYLPSPGFAEKYLSGISGFIKVQ 121

QY 281 LAEKQMPVRCLYKAGRAKESQGWYFTLENNLGEQDVCVFELLRTDFVLKVTAFRVNEY 340
    |||||
Db 122 LAEKQMPVRCLYKAGRAKESQGWYFTLENNLGEQDVCVFELLRTDFVLKVTAFRVNEY 181

QY 341 V 341
    |||||
Db 182 V 182

RESULT 10
AAG58317
ID AAG58317 standard; protein; 224 AA.
XX
AC AAG58317;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 75264.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
```

PR	16-APR-1999,	99US-0129845P.	PR	19-JUL-1999,	99US-0144334P.
PR	19-APR-1999,	99US-0130077P.	PR	19-JUL-1999,	99US-0144335P.
PR	21-APR-1999,	99US-0130449P.	PR	20-JUL-1999,	99US-0144352P.
PR	23-APR-1999,	99US-0130510P.	PR	20-JUL-1999,	99US-0144632P.
PR	23-APR-1999,	99US-0130891P.	PR	20-JUL-1999,	99US-0144884P.
PR	28-APR-1999,	99US-0131449P.	PR	21-JUL-1999,	99US-0144814P.
PR	30-APR-1999,	99US-0132048P.	PR	21-JUL-1999,	99US-0145086P.
PR	30-APR-1999,	99US-0132407P.	PR	21-JUL-1999,	99US-0145088P.
PR	04-MAY-1999,	99US-0132484P.	PR	22-JUL-1999,	99US-0145085P.
PR	05-MAY-1999,	99US-0132485P.	PR	22-JUL-1999,	99US-0145087P.
PR	06-MAY-1999,	99US-0132486P.	PR	22-JUL-1999,	99US-0145089P.
PR	08-MAY-1999,	99US-0132487P.	PR	22-JUL-1999,	99US-0145192P.
PR	07-MAY-1999,	99US-0132863P.	PR	23-JUL-1999,	99US-0145145P.
PR	11-MAY-1999,	99US-0134256P.	PR	23-JUL-1999,	99US-0145218P.
PR	14-MAY-1999,	99US-0134218P.	PR	23-JUL-1999,	99US-0145224P.
PR	14-MAY-1999,	99US-0134219P.	PR	26-JUL-1999,	99US-0145276P.
PR	14-MAY-1999,	99US-0134221P.	PR	27-JUL-1999,	99US-0145913P.
PR	14-MAY-1999,	99US-0134321P.	PR	27-JUL-1999,	99US-0145918P.
PR	14-MAY-1999,	99US-0134370P.	PR	27-JUL-1999,	99US-0145919P.
PR	18-MAY-1999,	99US-0134768P.	PR	28-JUL-1999,	99US-0145951P.
PR	19-MAY-1999,	99US-0134941P.	PR	02-AUG-1999,	99US-0146386P.
PR	20-MAY-1999,	99US-0135124P.	PR	02-AUG-1999,	99US-0146388P.
PR	21-MAY-1999,	99US-0135353P.	PR	02-AUG-1999,	99US-0146389P.
PR	24-MAY-1999,	99US-0135629P.	PR	03-AUG-1999,	99US-0147038P.
PR	25-MAY-1999,	99US-0136021P.	PR	03-AUG-1999,	99US-0147038P.
PR	27-MAY-1999,	99US-0136392P.	PR	04-AUG-1999,	99US-0147204P.
PR	28-MAY-1999,	99US-0136782P.	PR	04-AUG-1999,	99US-0147302P.
PR	01-JUN-1999,	99US-0137222P.	PR	05-AUG-1999,	99US-0147192P.
PR	03-JUN-1999,	99US-0137528P.	PR	05-AUG-1999,	99US-0147260P.
PR	04-JUN-1999,	99US-0137502P.	PR	06-AUG-1999,	99US-0147303P.
PR	07-JUN-1999,	99US-0137724P.	PR	06-AUG-1999,	99US-0147416P.
PR	08-JUN-1999,	99US-0138094P.	PR	09-AUG-1999,	99US-0147493P.
PR	10-JUN-1999,	99US-0138540P.	PR	09-AUG-1999,	99US-0147935P.
PR	10-JUN-1999,	99US-0138947P.	PR	10-AUG-1999,	99US-0148171P.
PR	14-JUN-1999,	99US-0139119P.	PR	11-AUG-1999,	99US-0148319P.
PR	16-JUN-1999,	99US-0139452P.	PR	12-AUG-1999,	99US-0148341P.
PR	16-JUN-1999,	99US-0139453P.	PR	13-AUG-1999,	99US-0148565P.
PR	17-JUN-1999,	99US-0139492P.	PR	13-AUG-1999,	99US-0148684P.
PR	18-JUN-1999,	99US-0139454P.	PR	16-AUG-1999,	99US-0149368P.
PR	18-JUN-1999,	99US-0139455P.	PR	17-AUG-1999,	99US-0149175P.
PR	18-JUN-1999,	99US-0139456P.	PR	18-AUG-1999,	99US-0149426P.
PR	18-JUN-1999,	99US-0139457P.	PR	20-AUG-1999,	99US-0149722P.
PR	18-JUN-1999,	99US-0139458P.	PR	20-AUG-1999,	99US-0149723P.
PR	18-JUN-1999,	99US-0139459P.	PR	20-AUG-1999,	99US-0149929P.
PR	18-JUN-1999,	99US-0139460P.	PR	23-AUG-1999,	99US-0149902P.
PR	18-JUN-1999,	99US-0139461P.	PR	23-AUG-1999,	99US-0149930P.
PR	18-JUN-1999,	99US-0139462P.	PR	23-AUG-1999,	99US-0149930P.
PR	18-JUN-1999,	99US-0139463P.	PR	25-AUG-1999,	99US-0150566P.
PR	18-JUN-1999,	99US-0139453P.	PR	26-AUG-1999,	99US-0150884P.
PR	23-JUN-1999,	99US-0140354P.	PR	27-AUG-1999,	99US-0151065P.
PR	24-JUN-1999,	99US-0140695P.	PR	27-AUG-1999,	99US-0151066P.
PR	18-JUN-1999,	99US-0139763P.	PR	27-AUG-1999,	99US-0151080P.
PR	21-JUN-1999,	99US-01			

```
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 52.7%; Score 946; DB 3; Length 224;
Best Local Similarity 81.2%; Pred. No. 2.6e-94;
Matches 190; Conservative 14; Mismatches 18; Indels 12; Gaps 5;

QY 110 MSAHNP-KRRLFDLDEDAEVIFFPSVSPLESTVPANKY-ASSAQTLTGTP 167
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      1 MDSAQNFKRRLFDPELDAKVIYPSN-----PESTPVMKGGSTAIQSPFK-E 53

QY 168 VKAEETPTPKPKKGRKKKADPEINSSAPRDDDPENRSKFYESASARKTVAER 227
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      54 SKABE---TPKVLKKGKKKKNPNEVNSTPGDDSDENRSKFYESASARKTVAER 110

QY 228 ERAINAAKTFEPTNPFRVLRPSLYRGICIMYLPSPGFAKYLSGISGFIKVQLAEKQWP 287
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      111 ERAVNAAKTFEPTNPFRVLRPSLYRGICIMYLPSPGFAKYLSGISGFIKVQLAEKQWP 170

QY 288 VRCLYKAGRAKFSQGWYFTLENNLGGDVCFVPELLRTDFVLKVTAPRVNEYV 341
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      171 VRCLYKAGRAKFSQGWYFTLENNLGGDVCFVPELLRTDFVLKVTAPRVNEYV 224

RESULT 11
AAG61371
ID AAG61371 standard; protein; 224 AA.
XX
AC AAG61371;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 79595.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000BP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 21-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142053P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
```

PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159233P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159234P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145919P.	PR 26-OCT-1999;	99US-0161359P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0145388P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161920P.
PR 03-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161922P.
PR 04-AUG-1999;	99US-0147204P.	PR 28-OCT-1999;	99US-0161993P.
PR 05-AUG-1999;	99US-0147302P.	PR 29-OCT-1999;	99US-0162142P.
PR 05-AUG-1999;	99US-0147192P.		
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147493P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		
PR 13-AUG-1999;	99US-0148684P.		
PR 16-AUG-1999;	99US-0149368P.		
PR 17-AUG-1999;	99US-0149175P.		
PR 18-AUG-1999;	99US-0149426P.		
PR 20-AUG-1999;	99US-0149722P.		
PR 20-AUG-1999;	99US-0149723P.		
PR 20-AUG-1999;	99US-0149929P.		
PR 23-AUG-1999;	99US-0149902P.		
PR 23-AUG-1999;	99US-0149930P.		
PR 25-AUG-1999;	99US-0150566P.		
PR 26-AUG-1999;	99US-0150884P.		
PR 27-AUG-1999;	99US-0151065P.		
PR 27-AUG-1999;	99US-0151066P.		
PR 27-AUG-1999;	99US-0151080P.		
PR 30-AUG-1999;	99US-0151303P.		
PR 31-AUG-1999;	99US-0151438P.		
PR 01-SEP-1999;	99US-0151930P.		
PR 07-SEP-1999;	99US-0152363P.		
PR 10-SEP-1999;	99US-0153070P.		
PR 13-SEP-1999;	99US-0153758P.		
PR 15-SEP-1999;	99US-0154018P.		
PR 16-SEP-1999;	99US-0154039P.		
PR 20-SEP-1999;	99US-0154779P.		
PR 22-SEP-1999;	99US-0155139P.		
PR 23-SEP-1999;	99US-0155486P.		
PR 24-SEP-1999;	99US-0156599P.		
PR 28-SEP-1999;	99US-0156458P.		
PR 29-SEP-1999;	99US-0156596P.		
PR 04-OCT-1999;	99US-0157117P.		
PR 05-OCT-1999;	99US-0157753P.		
PR 06-OCT-1999;	99US-0157865P.		
PR 07-OCT-1999;	99US-0158029P.		
<hr/>			
Query Match 52.7%; Score 946; DB 3; Length 224;			
Best Local Similarity 81.2%; Pred. No. 2.6e-84;			
Matches 190; Conservative 14; Mismatches 18; Indels 12; Gaps 5;			
QY 110	MDSAHNHF-KEARLFEDLEDEDAEVIFFSPYVPSPLPESTVPANKY-ASSAIQTLFTGP 167		
DB 1	MDSAQNFNKEARLFEDPELDAKVIYPSN-----PESTEPVKNKGYSSTAIQSFFK-E 53		
QY 168	VKAEEPTPTKPKKRGKKKNADPEINSAPRDDDPENESKFYESASAKRTVTAAER 227		
DB 54	SKAEE---TPKVLKRGKKKNPNPEEVNSTPGDDSENKSKFYESASAKRTVTAAER 110		
QY 228	BRAINAAKTPEPTNPFPRVLRPSLYRGCMYLPSPGFAEKYLSGISGFIKVQLAEQWP 287		
DB 111	ERAVNAKTPEPTNPFPRVLRPSLYRGCMYLPSPGFAEKYLSGISGFIKVQLAEQWP 170		
QY 288	VRCLYKAGRAKFSQGWYEFTLENNLGGDVCVFELLRTDRDVLKVTAPRVNEYV 341		
DB 171	VRCLYKAGRAKFSQGWYEFTLENNLGGDVCVFELLRTDRDVLKVTAPRVNEYV 224		
RESULT 12			
AAG61370			
ID	AAG61370 standard; protein; 226 AA.		
XX	AAG61370;		
XX			
DT	18-OCT-2000 (first entry)		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 79594.		
XX			
KW	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-00301439.		

XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123348P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153707P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.



```
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.

PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156598P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 52.7%; Score 946; DB 3; Length 226;
Best Local Similarity 81.2%; Pred. No. 2.7e-84;
Matches 190; Conservative 14; Mismatches 18; Indels 12; Gaps 5;

QY 110 MDSAHNF-KCARLEDEDEDAVIFPSSVYPSLPSTVTANKY-ASSAIQLFTGP 167
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3 MDSAQNQFNKARLFEDPELKDAKVIYPSN-----PESTEPVNKGYGGSTAIOGFFK-E 55
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 168 VKAREPTPTPKPKKRKKKNADPEEINSSAPRDDDPENRSKFYESASARKRTVTAER 227
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 56 SKAEEL---TPKVLKCRGKKKKNNPFEVNSSTPGDDSENRSKFYESASARKRTVTAER 112
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 228 ERAINAAKTFTPTNPFRRVLRPSYLYRGCIYMLPSGFAEKYLSGISGFIKVLAEKQWP 287
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 113 ERAVNAAKTFTPTNPFRRVLRPSYLYRGCIYMLPSGFAEKYLSGISGFIKVLAEKQWP 172
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 288 VRCLYKAGRAKFSQGWYEFTLENNIGEDVCVFELLRTDRFVLKTAFRVNEYV 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 173 VRCLYKAGRAKFSQGWYEFTLENNIGEDVCVFELLRTDRFVLKTAFRVNEYV 226
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
AAGI7136
ID AAGI7136 standard; protein; 224 AA.
XX
AC AAGI7136;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18043.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
```

hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.

XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0128264P.  
PR 29-MAR-1999; 99US-0128785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128714P.  
PR 08-APR-1999; 99US-0128845P.  
PR 16-APR-1999; 99US-0130077P.  
PR 19-APR-1999; 99US-0130449P.  
PR 21-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134258P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147322P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149358P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.

PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161320P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 52.3%; Score 940; DB 3; Length 224;		
Best Local Similarity 80.8%; Pred. No. 1e-83;		
Matches 189; Conservative 14; Mismatches 19; Indels 12; Gaps 5;		
Qy	110	MDSAHNR-KRRLPELDEDAEVIFFSSVYPSLPPESTVPANKGY-ASSAIQTLFTGP 167
Db	1	MDSAQNQFNKRLRFDPELKDVKIYPSN-----PEFEPNKGYGSGTAIQSFFK-E 53
Qy	168	VKAEEPTTPKPKKRGKRRKNADPEEINSSAPRDDDPENRSKSFYASARKRTVTABER 227
Db	54	SKAEE--TPKVLKKRKKKNPPEEVNNSTPGGDDSENRSKFYASARKRTVTABER 110
Qy	228	ERAINAAKTFFETNFFRVLPSPYLRCIMYLPSPGFAKYLSGISGFIKVLAKQWP 287
Db	111	ERAVNAAKTFFETNFFRVLPSPYLRCIMYLPSPGFAKYLSGISGFIKVLAKQWP 170
Qy	288	VRCLYKAGRAKFSQGYEFTLNNIGEGDVCVFELLRTDFVLKVTAFRVNEYV 341
Db	171	VRCLYKAGRAKFSQGYEFTLNNIGEGDVCVFELLRTDFVLKVTAFRVNEYV 224
RESULT 15		
ID	AAG17135	
XX	AAG17135	standard; protein; 226 AA.
AC	AAG17135;	
XX		

DT	17-OCT-2000	(first entry)
XX		
DE	Arabidopsis	thaliana protein fragment SEQ ID NO: 18042.
XX		
KW	Protein identification;	signal transduction pathway; metabolic pathway;
KW	hybridisation assay;	genetic mapping; gene expression control; promoter;
KW	termination sequence.	
XX		
OS	Arabidopsis	thaliana.
XX		
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000;	2000EP-00301439.
XX		
PR	25-FEB-1999;	99US-0121825P.
PR	05-MAR-1999;	99US-0123180P.
PR	09-MAR-1999;	99US-0123548P.
PR	23-MAR-1999;	99US-0125788P.
PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.



Job time : 82 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 06:07:48 ; Search time 25 Seconds  
(without alignments)  
904.578 Million cell updates/sec

Title: US-10-088-187A-11  
Perfect score: 1796  
Sequence: 1 MRPFPFKLIFSSTIQEKRL.....LLRTRDFVLKVTAFRVNEYV 341

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/prodata/1/1aa/6A\_COMB.pep:\*
- 4: /cgn2\_6/prodata/1/1aa/6B\_COMB.pep:\*
- 5: /cgn2\_6/prodata/1/1aa/ECTUS\_COMB.pep:\*
- 6: /cgn2\_6/prodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97.5	5.4	1178	4	US-09-543-681A-7819
2	97.5	5.4	8991	4	US-08-714-741-32
3	97	5.4	399	4	US-09-107-532A-6398
4	95.5	5.3	834	3	US-08-539-205A-6
5	95.5	5.3	834	4	US-09-392-163A-6
6	94.5	5.3	497	4	US-09-248-796A-15909
7	94.5	5.3	693	4	US-09-248-796A-26890
8	93.5	5.2	274	4	US-09-489-039A-10361
9	92.5	5.2	194	4	US-08-529-055-64
10	92.5	5.2	496	3	US-08-740-223A-16
11	92.5	5.2	496	4	US-09-709-188-16
12	92	5.1	748	4	US-09-270-767-44371
13	91.5	5.1	1736	4	US-09-919-497-98
14	90	5.0	1061	4	US-09-762-481B-2
15	88.5	4.9	286	3	US-08-960-507-20
16	88.5	4.9	286	4	US-09-136-801-20
17	88.5	4.9	286	4	US-09-202-088A-20
18	88.5	4.9	334	4	US-09-248-796A-21169
19	88.5	4.9	480	3	US-08-740-223A-8
20	88.5	4.9	480	4	US-09-709-188-8
21	88.5	4.9	496	1	US-08-373-579-6
22	88.5	4.9	496	2	US-08-418-595-6
23	88.5	4.9	496	2	US-08-665-926-6
24	88.5	4.9	496	3	US-09-162-437-6
25	88.5	4.9	496	3	US-08-740-223A-6
26	88.5	4.9	496	3	US-09-351-457-4
27	88.5	4.9	496	3	US-09-561-500-4

28	88.5	4.9	496	3	US-09-561-108-4	Sequence 4, Appl
29	88.5	4.9	496	4	US-09-351-543-4	Sequence 4, Appl
30	88.5	4.9	496	4	US-09-561-526-4	Sequence 4, Appl
31	88.5	4.9	496	4	US-09-202-491-8	Sequence 8, Appl
32	88.5	4.9	496	4	US-08-817-318-6	Sequence 6, Appl
33	88.5	4.9	496	4	US-09-709-188-6	Sequence 6, Appl
34	88.5	4.9	496	4	US-09-561-499-4	Sequence 4, Appl
35	88.5	4.9	496	4	US-09-442-717-6	Sequence 6, Appl
36	88.5	4.9	496	4	US-09-689-020-6	Sequence 6, Appl
37	88.5	4.9	496	4	US-09-998-831-4	Sequence 4, Appl
38	88.5	4.9	496	4	US-09-561-005-4	Sequence 4, Appl
39	88.5	4.9	496	4	US-09-819-386-4	Sequence 4, Appl
40	88.5	4.9	498	3	US-08-740-223A-20	Sequence 20, Appl
41	88.5	4.9	498	4	US-09-709-188-20	Sequence 20, Appl
42	88.5	4.9	499	3	US-08-740-223A-24	Sequence 24, Appl
43	88.5	4.9	499	4	US-09-709-188-24	Sequence 24, Appl
44	87.5	4.9	389	4	US-09-248-796A-17225	Sequence 17225, A
45	87.5	4.9	432	4	US-09-248-796A-18913	Sequence 18913, A

ALIGNMENTS

RESULT 1

US-09-543-681A-7819  
; Sequence 7819, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543, 681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 7819  
; LENGTH: 1178  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-7819

Query Match		5.4%	Score 97.5;	DB 4;	Length 1178;
Best Local Similarity		21.7%	Pred. No. 0.47;		
Matches		53;	Conservative	40;	Mismatches 98;
				Indels	53;
				Gaps	10;
QY	15	IQEKRLR-----	VDPKFVSKPKDELSVAVALTPDGHVVRVGLRKADNKKIWFQDGM	65	
Db	432	IEEELKENTHEVHAIVPQIASYLLNEKRKAVT-----	DIELRQANVKV-----	476	
QY	66	QEFVDYRIGIRGYLLIFRYEGNSAFSVIFNLHSEINHYHSTGLMDSAHNHFKEARLFED	125		
Db	477	-VIVPNDQMTPHFVIRKGEIISLSYNLA-----	QYHETQLNEADEAVPERKAPQ	531	
QY	126	-----LEDEDARVIFPSSVYVSP-----	LPSTVPANKGYASSAIQTFTG-----	PV	168
Db	532	PAISAFALQDMSADTNNKKSAPTAKQVKSETTENGPF-PSAITKFSNLFASQPEPV	590		
QY	169	KAEPTPTPKIPKGRKKKNADPEEINSSAPRDDDPNRSKFYSASARKKTVTABERE	228		
Db	591	VEEKQESKKNENRRRRNNRRNRDND---RNRDHE---	RNRDNTSENR	644	
QY	229	RAIN 232			
Db	645	KNRN 648			

RESULT 2

US-08-714-741-32  
; Sequence 32, Application US/08714741  
; Patent No. 6500613  
; GENERAL INFORMATION:

APPLICANT: Briles, David E.  
APPLICANT: McDaniel, Larry S.  
APPLICANT: Swiatlo, Edwin  
APPLICANT: Yother, Janet  
APPLICANT: Crain, Marilyn J.  
APPLICANT: Hollingshead, Susan  
APPLICANT: Tart, Rebecca  
APPLICANT: Brooks-Walter, Alexis  
TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,  
TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,  
TITLE OF INVENTION: PORTIONS AND PRODUCTS  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/714,741  
FILING DATE: 16-SEP-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Prommer Esq., William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2460  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8991 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: amino acid  
US-08-714-741-32

Query Match 5.4%; Score 97.5; DB 4; Length 8991;  
Best Local Similarity 24.4%; Pred. No. 12;  
Matches 38; Conservative 17; Mismatches 72; Indels 29; Gaps 4;  
Qy 125 DLEDEDAEVFPSSVPSPLPESTVPANKGYASSAIQTLLFTGVPKAEETPTPKIPKRG 184  
Db DEETPAPAPQEQPAPAPKPEQAPA-----PKPEQAPAPK-PEQPA 5720  
Qy 185 RKKNADPEEINSAPRDDDPENRSKFYESSASARKRTVTAEERERAINAAKTPEPTNPF 244  
Db 5721 PAPPEQF---APAPKPEQAPKPEKPAEPTQPEKPAETKTRVRALKVAE-----F 5768  
Qy 245 RVULRPSYLRGCMYLPSPGFAEYLSGIGFIKVQ 280  
Db 5769 GVQLRDAGGNNVNGAYFKGLETTAEYXGLGKAE 5804

RESULT 3  
US-09-107-532A-6398  
Sequence 6398, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street

CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD/ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION/DOCKET NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 6398:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 399 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...399  
SEQUENCE DESCRIPTION: SEQ ID NO: 6398:  
US-09-107-532A-6398

Query Match 5.4%; Score 97; DB 4; Length 399;  
Best Local Similarity 28.9%; Pred. No. 0.095;  
Matches 37; Conservative 14; Mismatches 71; Indels 6; Gaps 5;  
Qy 118 KRALFEDLEDEDAEVFPSSVPSPLPESTVPANKGYASSAIQTLLFTGVPKAEETPTP 177  
Db 75 KQPTAPEMAFAEQTEPLITPEYLADSTAEPDPPAPASMRGAVDMPLLDAP1-AEETTPAP 133  
Qy 178 --KTPKKGRKKKNADPEEINSAPRDDDPENRS-KFYESSASARKRTVTAEERERAINAA 234  
Db 134 VRKQPKRKGKGGKKQKTKFTEDDARPGSEKTAIRGVEGAPAPKRR-TSEGGSR-LQFD 191  
Qy 235 KTFEPTNP 242  
Db 192 KEEEPADP 199

RESULT 4  
US-08-539-205A-6  
Sequence 6, Application US/08539205A  
Patent No. 6001619  
GENERAL INFORMATION:  
APPLICANT: Beach, David H.  
APPLICANT: Caligiuri, Maureen  
APPLICANT: Nefsky, Bradley  
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/539,205A  
FILING DATE: 04-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: CSV-005.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 834 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-539-205A-6

Query Match 5.3%; Score 95.5; DB 3; Length 834;  
Best Local Similarity 21.0%; Pred. No. 0.45;  
Matches 66; Conservative 32; Mismatches 109; Indels 107; Gaps 15;

QY	1	MPRPFFKLIFSTTQEKRLVPDKFVSKFDLSVAVALTVPDG-----H 46
DB	1	MERP:-----TKDFLLR-PRSHKSRVKGFLKXWAYMPKNGQDENSDQDDMEH 51
QY	47	VVRVGLRKADNKKWFQD-----GWOEFVDYRSIRIGYLLIFRYEGNSAFSVYIFNL 97
DB	52	GWEV-VDSNDSASQHQEELPPPLPGWEEKVDNLG-----RTYYVNH 93
QY	98	SHSEINYHSTGLMD-----SAHHFK-RARLPEDLEDEDAEVIFFSSVYP 141
DB	94	NNRTQWHRPRLMDVSSSDNNIRQINQEAHRRFRSRHISEDLEPE-----PSEGDD 147
QY	142	SPLPESTVPANKGYASSAIQTLFTGPVKAEBEPTPKPKKGRKKGNADPEEINSAPR 201
DB	148	VPEPWETISEEVNIAGDSLGLAL-----PPPVSPGSR-----TSPQELSEELSR 192
QY	202	-----DDDPENRSKPYE-SASARKRTVTAERERAINAAKTFTPTNPFVRVLRPSVLY 254
DB	193	RLQITPDSNGEQFSLLIQREPSSRLRSCSVTD---AAVEQGHLP-----PSVAY 239
QY	255	RGCIMYLPSPGPAEK 268
DB	240	VHTTGLPSGWEER 253

RESULT 5  
US-09-392-163A-6  
Sequence 6, Application US/09392163A  
Patent No. 6503742  
GENERAL INFORMATION:  
APPLICANT: Beach, David H.  
APPLICANT: Caligiuri, Maureen  
APPLICANT: Netsky, Bradley  
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/392,163A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/539,205  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: CSV-005.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 834 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-392-163A-6

Query Match 5.3%; Score 95.5; DB 4; Length 834;  
Best Local Similarity 21.0%; Pred. No. 0.45;  
Matches 66; Conservative 32; Mismatches 109; Indels 107; Gaps 15;

QY	1	MPRPFFKLIFSTTQEKRLVPDKFVSKFDLSVAVALTVPDG-----H 46
DB	1	MERP:-----TKDFLLR-PRSHKSRVKGFLKXWAYMPKNGQDENSDQDDMEH 51
QY	47	VVRVGLRKADNKKWFQD-----GWOEFVDYRSIRIGYLLIFRYEGNSAFSVYIFNL 97
DB	52	GWEV-VDSNDSASQHQEELPPPLPGWEEKVDNLG-----RTYYVNH 93
QY	98	SHSEINYHSTGLMD-----SAHHFK-RARLPEDLEDEDAEVIFFSSVYP 141
DB	94	NNRTQWHRPRLMDVSSSDNNIRQINQEAHRRFRSRHISEDLEPE-----PSEGDD 147
QY	142	SPLPESTVPANKGYASSAIQTLFTGPVKAEBEPTPKPKKGRKKGNADPEEINSAPR 201
DB	148	VPEPWETISEEVNIAGDSLGLAL-----PPPVSPGSR-----TSPQELSEELSR 192
QY	202	-----DDDPENRSKPYE-SASARKRTVTAERERAINAAKTFTPTNPFVRVLRPSVLY 254
DB	193	RLQITPDSNGEQFSLLIQREPSSRLRSCSVTD---AAVEQGHLP-----PSVAY 239
QY	255	RGCIMYLPSPGPAEK 268
DB	240	VHTTGLPSGWEER 253

RESULT 6  
US-09-248-796A-15909  
Sequence 15909, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 15909  
LENGTH: 497  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-15909

```
Query Match      5.3%; Score 94.5; DB 4; Length 497;
Best Local Similarity   21.0%; Pred. No. 0.25;
Matches    51; Conservative     40; Mismatches       95; Indels     57; Gaps        9;

y          9 LIFSQTIOEKRLRVPDKFVSKFKDLSVAVALTPDGHWVRVG---LRKADNKIWFQDGM 65
|::||::||::||::||::||::||::||::||::||::||::||::||:
b         260 LLNSNDDEPIRKQINFGMRDPKNFLNLANPFTTSSSERLDENYVTANN----- 311
|::||::||::||::||::||::||::||::||::||::||::||::||:
y          66 QEFVDYSIRIGYLIFRYEGNSAFSVFIYNLSHSEINYHSTGLMDSAHNHFKARLPED 125
|::||::||::||::||::||::||::||::||::||::||::||::||:
b         312 -----DYHFIFRNPGDPILFDLNQHYTEVOFIQITSDKSSELHSK----- 353
|::||::||::||::||::||::||::||::||::||::||::||::||:
y          126 LEDDAEVFPSPSVPSPLESTTVPAANKGYASSAQTLTFTPVKAEETPT----- 176
|::||::||::||::||::||::||::||::||::||::||::||::||:
b         354 --EIDAEELIKGUTIEAPIOK-IOSNE----TATTASTTKTTKTTPSLTRLTTTTV 406
|::||::||::||::||::||::||::||::||::||::||::||::||:
y          177 ---PKIPKKRGKKGNAPDEINSRAPRDDPENRSKFYESASARKRTVTAAERERAINA 233
|::||::||::||::||::||::||::||::||::||::||::||::||:
b         407 FGKPSITKKG-KHLINKDVVSQGTINGD-----PDGVTYGOPEPERDRRA-NY 457
|::||::||::||::||::||::||::||::||::||::||::||::||:
y           234 AKT 236
|::||::||::||::||::||::||::||::||::||::||::||::||:
b          458 AET 460

RESULT 7
S-09-248-796A-26890
Sequence 26890, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
Applicant: Keith Weinstein et al
TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO CANDII
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 26890
LENGTH: 693
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-26890

Query Match      5.3%; Score 94.5; DB 4; Length 693;
Best Local Similarity   20.5%; Pred. No. 0.43;
Matches    54; Conservative     39; Mismatches       116; Indels     55; Gaps        8;

yy         10 IPFS-----TIQEKLRVDP-----KFYSKFKDLSVAALTPDGHWVRVG 51
|::||::||::||::||::||::||::||::||::||::||::||::||:
bb        166 IFSSIILLIWTRQYTIIDQRCKKVFNVNIALQQFWLFVIIITFTDSLQOSSMHNLIS 225
|::||::||::||::||::||::||::||::||::||::||::||::||:
yy         52 LRKADNKIWFODGWQOFEVDYSIRIGYLLIFRYEGNSAFSVFIYNLSHSEINYHSTGLMD 111
|::||::||::||::||::||::||::||::||::||::||::||::||:
bb        226 WETQLKNWYQDLISHYFLPNASNQI--LLDHLQHNRQFDOLKMDDL I----- 270
|::||::||::||::||::||::||::||::||::||::||::||::||:
yy        112 SAHNHFKAARLPEDLEDADAIVFPSVSYPL-PESTVPANKGYASSAQTLFTGPVKA 170
|::||::||::||::||::||::||::||::||::||::||::||::||:
db        271 SLONELEBKSLSNSDRESSFTPPFLNNSTPLLSPNPNSNYVSFFPKAPPSS 330
|::||::||::||::||::||::||::||::||::||::||::||::||:
yy        171 EEPTPTPKIPKRGRKKGNADOPE-EINSNAPRDOPDENRSKFYESASARKRTVTAERER 2299
|::||::||::||::||::||::||::||::||::||::||::||::||:
db        331 ELPVSSPVPRKSGSTWKSSKKQLNTIIIEDD-----STLYDSKE----- 372
|::||::||::||::||::||::||::||::||::||::||::||::||:
yy        230 AINAAKTEPTNPFRVVLRPSYL 253
|::||::||::||::||::||::||::||::||::||::||::||::||:
db        373 --NTFTNFEDIIPGIGITTREP TL 394
|::||::||::||::||::||::||::||::||::||::||::||::||:
```

## RESULT 8

```

US-09-489-039A-10361
; Sequence 10361, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breston et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10361
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10361

Query Match      5.2%; Score 93.5; DB 4; Length 274;
Best Local Similarity 24.4%; Pred. No. 0.12;
Matches 51; Conservative 20; Mismatches 65; Indels 73; Gaps 9;

Qy      125 DLEDED-AEVFPSSVVPSPLESTV---PANKGYASSAIQTLTFTGVPKAEETPTPK-- 178
Db      83 DLEPPPAQPVWPEPEPEPEPEVVPPEPKE-----APVVIHKPEKPKPK 130

Qy      179 -IPKRGK--K-----KNADPBEINSSAPRDDDDPENRSKFYESASARKETVT 223
Db      131 PKPKPEKPEKQVKREKVAEPRPASPENNNTARTAPST-----STAAAKPTVT 184

Qy      224 AERERAINAAKTFTPTNPPFRVVLRSYVYRGCMYLPDGFABKYLSGISGFI----- 277
Db      185 APSGPRALSR-----VQSPYPARAQAALRIEGTVRVKFDVSPDGRIDNLQIL 230

Qy      278 -----KVQLAEKQWVRCCLYKAGR 296
Db      231 SAQPANMFREVKSAMRRW-----RYQOGR 255

RESULT 9
US-08-529-055-64
; Sequence 64, Application US/08529055
; Patent No. 6592876
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Votter, Janet
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: Pneumococcal Genes, Portions
; TITLE OF INVENTION: Thereof, Expression Products
; TITLE OF INVENTION: Therefrom, and Uses of Such Genes,
; TITLE OF INVENTION: Portions and Products
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,055
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.

```

```

; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2400
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-529-055-64

```

Query Match	5.2%;	Score 92.5;	DB 4;	Length 194;
Best Local Similarity	28.1%;	Pred. No. 0.092;		
Matches 36;	Conservative 17;	Mismatches 56;	Indels 19;	Gaps 4;
Qy	118	KRALFLDLEDEDAEVIFFSPSVYPSLPPESTVPANKGVASSAIQTLFTGPKVKE---	PT 174	
Db	81	KKALEGTETADLLKANVNEPKGAPAPAPETPAP-----EAPAQ-----	PKAPETTPAPA 130	
Qy	175	PTPKIPKKRKKKNADPPEEINSSAPRDDDDPENRSKFYESASARKTVTAEERERAINAA	234	
Db	131	PKPEKPAEQPKPEKPAQQAEDYARSEEBEYNLTQQQPAPAKQPKPEQPAKPEKPA	--- 186	
Qy	235	KTPEPTNP	242	
Db	187	--EPTQP	192	

```

RESULT 10
; US-08-740-223A-16
; Sequence 16, Application US/08740223A
; Patent No. 6265564
; GENERAL INFORMATION:
; APPLICANT: Davis, et al.
; TITLE OF INVENTION: Expressed Ligand - Vascular
; TITLE OF INVENTION: Intercellular Signalling Molecule
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.

```

STREET: 777 Old Saw Mill Road  
CITY: Tarrytown  
STATE: NY  
COUNTRY: USA  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/740.223A  
FILING DATE: 25-OCT-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 60/022/999  
FILING DATE: 02-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Cobert, Robert J  
REGISTRATION NUMBER: 36,108  
REFERENCE/DOCKET NUMBER: REG 333  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 914-345-7400  
TELEFAX: 914-345-7721  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 496 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein

```

; ; FEATURE:
; ; NAME/key: htl2
; ; LOCATION: 1...496
; ; OTHER INFORMATION: human TIE-2 ligand 2
US-08-740-223A-16

Query Match          5 24; Score 92.5; DB 3; Length 496;
Best Local Similarity 29.3%; Pred. No. 0.42; Gaps 4;
Matches 27; Conservative 16; Mismatches 26; Indels 23; Gaps 4;

Qy      45   GHVVRVLGRKADNNKIWFQDGWQ-----EFVDYSIRIGYLL---IFRY 84
         ||| : | : | : | : | : | : | : | : | : | : | : | : | :
Db      319  GGGWTIIQRREDGSLDFQKGWKVGFGSPSGEYWLGNFEFSQITNQRYVLKHLKD 378

Qy      85   EGNSAFVY--IFNLHSHSEINY--HSTGLMDSA 113
         ||| : | : | : | : | : | : | : | : | : | : | : | : |
Db      379  EGNEAYSLYDHFYISGBELNYRIHLKGLTGTA 410

```

```

RESULT 11 .
US-09-709-188-16
; Sequence 16, Application US/09709188
; Patent No. 6441137
; GENERAL INFORMATION:
; APPLICANT: Davis et al.
; TITLE OF INVENTION: Expressed Ligand - Vascular Intercellular Signaling Molecule
; FILE REFERENCE: REG 333-Z
; CURRENT APPLICATION NUMBER: US/09/709,188
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 08/740,223
; PRIOR FILING DATE: 1996-10-25
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-709-188-16

```

```

Query Match      5.2%; Score 92.5; DB 4; Length 496;
Best Local Similarity 29.3%; Pred. No. 0.42;
Matches 27; Conservative 16; Mismatches 26; Indels 23; Gaps 4;

Qy 45 GHWVRVGLRKADNKIWFQGWQ-----EFVDYSIRICYLL---IFRY 84
      ||| : : : ||| : : : ||| : : : ||| : : :
Db 319 GCGWTIIQRREDGSLDFQXGWEKYVFGSPSGEYWLGNFEISQITNQRYVLKHLKDW 378
      ||| : : : ||| : : : ||| : : : ||| : : :

Qy 85 EGNSAFSVY-IFNLHSEINY--HSTGLMDSA 113
      ||| : : : ||| : : : ||| : : : ||| : : :
Db 379 EGNEAYSLYDHFYISGEELNRIHLKGLTGA 410
      ||| : : : ||| : : : ||| : : : ||| : : :

```

```

RESULT 12
US-09-270-767-44371
; Sequence 44371, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44371
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-44371

```

Query Match 5.1%; Score 92; DB 4; Length 748;  
Best Local Similarity 27.28; Pred. No. 0.91;  
Matches 44; Conservative 19; Mismatches 63; Indels 36; Gaps 7;

QY 98 SHSINHYSTGLMSAHNHFKARLFDLEDEDAEVI--FPSSVYPSLPST-----VPA 151  
Db 536 SSSILEKHLTGRKLLHH-----SAVNDDARVLLEFANSKOPPLAASSTTFVNA 588  
QY 152 NKGVASSAIQTLFTGPKVKAEBPTTPKPKKGRK---KNADPEEINSAPRDDDPEN 207  
Db 589 SVGPSSA-----GSVFASSTSCSKAPKGEYILPDNMDVDEIVIS-----S 633  
QY 208 RSKFYESASAKRTVTAE-----RERAINAKTPEPTNPPFR 245  
Db 634 SSSYTSSAAQVNTLQPDNFELXHQATTATTATYFNFGR 675

RESULT 13  
US-09-919-497-98  
; Sequence 98, Application US/09919497  
; Patent No. 6773883  
; GENERAL INFORMATION:  
; APPLICANT: Mutter, George L.  
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER  
; FILE REFERENCE: B0801/7225  
; CURRENT APPLICATION NUMBER: US/09/919,497  
; PRIOR FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/221,735  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 98  
; LENGTH: 1736  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-919-497-98

Query Match 5.1%; Score 91.5; DB 4; Length 1736;  
Best Local Similarity 19.4%; Pred. No. 4;  
Matches 63; Conservative 41; Mismatches 126; Indels 95; Gaps 14;

QY 44 DGHVVRGLRKA-----DNK1WFQD-----WQEFVDYRSIRIGYLLIFRYEGNSAFSVY 93  
Db 768 DG--WYGAKEAVQQOQNLVWVSEKAGDAGTSDLDLHDDRLSYL-----SAPGEYSWY 821  
QY 94 IPNLSHSEINHYSTGLMSAHNHFKARLFDLEDE-----DAEVIFFPS----- 138  
Db 822 STDSTRHS-DYEDT---DTEGAYTDQELDELINDEVTGTPPESAITRSSEPVREDSSGMH 877  
QY 139 -----VYP-----SPLPESTVPANKGYASSAI-OTL 163  
Db 878 HENQTYPPYSPQAQPIHRIDSPGPKPASQQAASSFPVYLSPETNPASTSAVNHV 937  
QY 164 FTGPKVKAEBPTTPKPKKGRKKKNADPEEINSAPRDDDPENRSKFYESASAKRTVT 223  
Db 938 NUTNVRLEPTPAPSTYSYQADSURTSTSTEAHMLRDQEP-SLSSHVDFTKVKRDPY 996  
QY 224 AEERERAINAK-----TPEPTNPPF-----RVVLRPSYLYRCIMYL 261  
Db 997 PEEMRQNHVLKQPAVSHGHRPDKEPNLTYPQLPYVEKQASRDLEQPTRYES--SSY 1054  
QY 262 PSGFAEKYLSGISGFIKVLAKQW 286  
Db 1055 TDQFSRNYEHLRYEDRVPMYEQW 1079

RESULT 14  
US-09-762-481B-2  
; Sequence 2, Application US/09762481B  
; Patent No. 6632639  
; GENERAL INFORMATION:  
; APPLICANT: DREYFUS, MARC  
; APPLICANT: LOPEZ, PASCAL  
; TITLE OF INVENTION: MUTANT E. COLI STRAINS, AND THEIR USE FOR PRODUCING  
; TITLE OF INVENTION: RECOMBINANT POLYPEPTIDES  
; FILE REFERENCE: US9898APCNR

QY 93 YIFNLHSEINHYST-----GLMDSAHNHFKARLFDL----- 126  
Db 450 YLLNEKRSAVNAITETRODGVRCVIVPNDQMETPHVHLVRKGBETPTLSYMLPKLHEEA 509  
QY 127 -----EDEDAEVIFPS-----SVYSPLPES-----TVPANKG 154  
Db 510 MALPSEEEFAERKPEQPALATFAMPDVPPTPAEPAAPVVPAPAKAATPAAPAQPG 569  
QY 155 YAS---SAIOTLFTGPKVKAEBPTT---PKPKKGRKKKNADPEEINSAPRDDDPEN 207  
Db 570 LLSRFFGALKALFSG---GEETKTEQAPAKAEAKPERQODRRKPRQ--NNRDRNERRDT 625  
QY 208 RSKFYESASAK-----RTVTAERERAINAKT 236  
Db 626 RSERTEGSDNRENNRRNRQAQQQTAEATRESRQQAQVET 663

RESULT 15  
US-08-960-507-20  
; Sequence 20, Application US/08960507  
; Patent No. 6057435  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; TITLE OF INVENTION: Tie Ligands  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/960,507  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dreger, Ginger R.  
; REGISTRATION NUMBER: 33,055  
; REFERENCE/DOCKET NUMBER: P1130p1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-3216  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 286 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-960-507-20

Query Match 4.9%; Score 88.5; DB 3; Length 286;  
Best Local Similarity 30.4%; Pred. No. 0.47;  
Matches 28; Conservative 14; Mismatches 27; Indels 23; Gaps 4;  
QY 45 GHVRYGLRKADNKIWFDGQWQ-----EFVDRYSIRIGYLL---IFRY 84  
Db 109 GGGWTIIQRREDSVDFQRTWKYKVGFGNPSGEYWLGNFVSQLTNQORYVLKHLKDW 168  
QY 85 EGNsafsvY-IFNLshSEINy--HSTGLMDSA 113  
Db 169 EGNEAYSLYEHFYLSSSEELNYRIHLKGLTCTA 200

Search completed: December 30, 2004, 06:15:51  
Job time : 27 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 06:10:48 ; Search time 367 Seconds  
(without alignments)  
334.243 Million cell updates/sec

Title: US-10-088-187A-11

Perfect score: 1796

Sequence: 1. MPRPFHKLIFSSTIQEKRL.....LLRTRDFVLKVTAFRNEYV 341

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	908.5	50.6	431	15	US-10-424-599-260947
2	554.5	30.9	188	15	US-10-424-599-240434
3	375.5	20.9	203	15	US-10-424-599-240433
4	329	18.3	344	17	US-10-739-930-9287
5	289.5	16.1	502	17	US-10-425-115-274348
6	280	15.6	375	15	US-10-425-114-63507
7	278	15.5	402	16	US-10-437-963-157411
8	266.5	14.8	181	15	US-10-424-599-166413
9	265	14.8	362	16	US-10-437-963-192114
10	251.5	14.0	750	16	US-10-437-963-178183
11	246	13.7	737	16	US-10-437-963-203869
12	241.5	13.4	322	16	US-10-437-963-152254
13	214	11.9	306	16	US-10-437-963-129669

14	210.5	11.7	164	15	US-10-424-599-206969	Sequence 206969,
15	204.5	11.4	396	17	US-10-425-115-314879	Sequence 314879,
16	197	11.0	519	16	US-10-437-963-187150	Sequence 187150,
17	195.5	10.9	1355	16	US-10-437-963-137649	Sequence 137649,
18	193.5	10.8	118	15	US-10-424-599-179398	Sequence 179398,
19	193	10.7	536	16	US-10-437-963-162720	Sequence 162720,
20	186	10.4	462	16	US-10-437-963-195577	Sequence 195577,
21	183	10.2	100	15	US-10-424-599-232448	Sequence 232448,
22	175.5	9.8	538	16	US-10-437-963-120554	Sequence 120554,
23	174	9.7	422	16	US-10-437-963-177245	Sequence 177245,
24	171.5	9.5	167	17	US-10-739-930-9695	Sequence 9695, Ap
25	170.5	9.5	391	16	US-10-437-963-138423	Sequence 138423,
26	166	9.2	270	16	US-10-437-963-156915	Sequence 156915,
27	165.5	9.2	181	15	US-10-424-599-241998	Sequence 241998,
28	162	9.0	259	15	US-10-424-599-241690	Sequence 241690,
29	161	9.0	352	15	US-10-425-114-42561	Sequence 42561, A
30	147	8.2	305	16	US-10-437-963-105662	Sequence 105662,
31	145.5	8.1	230	16	US-10-437-963-188750	Sequence 188750,
32	141	7.9	143	17	US-10-425-115-274347	Sequence 274347,
33	141	7.9	417	16	US-10-437-963-170458	Sequence 170458,
34	140.5	7.8	141	16	US-10-437-963-195574	Sequence 195574,
35	140.5	7.8	156	15	US-10-424-599-207209	Sequence 207209,
36	140	7.8	255	16	US-10-437-963-188901	Sequence 188901,
37	135.5	7.5	105	17	US-10-425-115-292035	Sequence 292035,
38	135.5	7.5	212	17	US-10-425-115-253856	Sequence 253856,
39	134	7.5	237	16	US-10-437-963-122652	Sequence 122652,
40	130	7.2	239	17	US-10-425-115-341488	Sequence 341488,
41	128.5	7.2	1070	16	US-10-437-963-167140	Sequence 167140,
42	127.5	7.1	226	17	US-10-425-115-358357	Sequence 358357,
43	126.5	7.0	156	17	US-10-425-115-221099	Sequence 221099,
44	123.5	6.9	148	15	US-10-424-599-173953	Sequence 173953,
45	122.5	6.8	87	16	US-10-767-701-54013	Sequence 54013, A

ALIGNMENTS

RESULT 1  
US-10-424-599-260947  
; Sequence 260947, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 260947  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_77658C.1.pep  
US-10-424-599-260947

Query Match	50.6%	Score	908.5	DB	15	Length	431
Best Local Similarity	49.7%	Pred. No.	7.1e-76				
Matches	198	Conservative	46	Mismatches	95	Indels	59
Gaps	8						
QY	1	MPRPFHKLIFSSTIQ-EKRLRVPDKFVSKFKDELSSVALVTPDGHVWRVGLRKADNKI	59				
Db	1	MPHPSFKLLLPSTVQPNQQLRPLDNFMKYGGLSPITVTLSPDGSVWVHVLGKADNKY	60				
QY	60	WFQDQWQFVDYRYSIRIGYLLIFRYEGNSASVYIFNLSHSEINHTSLGLMDSAHNFKR	119				
Db	61	CFLDGWKEFQVRSYIGVSLCVRTRVGKSVRTWIFNLATSGINYSQSVTRSSNEGLHFTN	120				
QY	120	A-RLPDEDEDAEVI---FPSSVPSPLP-----ESTVPANKGVASSAIQTLFTG-	166				

```
Db 121 CLKFEEMEGEDSIEISDSSPHSLSPSSIQNALAGSDVKMPGKSYNTPPALQNLWNGS 180
Qy 167 -----PVAEPTTPPKI-----PKRG 184
Db 181 KLSNWGEGGNAHSRSANSIDNLRTRDIGQFNAVEFKRSTELKLRASIEERMKKT 240
Qy 185 RKKNADPEEINSSAPRDDPENRSKFYESASARKRTVTAERERAINAAKTFFPTNPF 244
Db 241 RKGRSDGQE--PSAGHEEVEEMRFRFYESASARKRTVTAERERIVNEAKAFESNFC 298
Qy 245 RVVLRPSYLYRGCMYLPSPGFAEKYLSGISGFIKVQLAE-KQMPVRCLYKAGRAKFSQGW 303
Db 299 RVVLRPSYLYRGCMYLPSCFAEKHLNGVSGFIKQLISNGRQWVPVRCYKGRAKLSQGW 358
Qy 304 YFTELENNLGEQDVCVFELLRTDRFVLKVTAPRVNEYV 341
Db 359 FEFSLLENNLGEQDVCVFELLRMKEVVLQVTFPHVTEV 396
```

## RESULT 2

US-10-424-599-240434

; Sequence 240434, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 240434

; LENGTH: 188

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_59139C.1.pap

US-10-424-599-240434

```
Query Match 30.9%; Score 554.5; DB 15; Length 188;
Best Local Similarity 76.6%; Pred. No. 2.5e-43;
Matches 108; Conservative 14; Mismatches 14; Indels 5; Gaps 2;
```

```
Qy 204 DPENRS----KFYESASARKRTVTAERERAINAAKTFFPTNPFRRVLRPSYLYRGCM 259
Db 7 DSENVADKPRWRYELASARKSTVTAERERAINASKTFFPTNPFRCVLRPSYLYRGCM 66
```

```
Qy 260 YLPSCFAEKYLSGISGFIKVQLAE-KQMPVRCLYKAGRAKFSQGWYFTELENNLGEQDVC 318
Db 67 YLPSCFAEKNLNGVSGFIKQLISNGRQWSVRCYKGRAKLSQGWFEFTVENNLGEQDVC 126
```

```
Qy 319 VFELLRTDRFVLKVTAPRVNE 339
Db 127 VFELLRTREVLQVTVFVTE 147
```

## RESULT 3

US-10-424-599-240433

; Sequence 240433, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

```
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 240433
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_59138C.1.pap
US-10-424-599-240433
```

```
Query Match 20.9%; Score 375.5; DB 15; Length 203;
Best Local Similarity 68.5%; Pred. No. 1.4e-26;
Matches 74; Conservative 14; Mismatches 19; Indels 1; Gaps 1;
```

```
Qy 182 KRGRKKKNADPEEINSSAPRDDPENRSKFYESASARKRTVTAERERAINAAKTFFPTN 241
Db 96 KTAKKKRRKSEPYGEEPSGENEEAEEMRYRFYESASARKRTVTAERERAINASKTFFPTN 155
Qy 242 PFFRVLRPSYLYRGCMYLPSPGFAEKYLSGISGFIKVQLAE-KQMPV 288
Db 156 PFCVLRPSYLYRGCMYLPSTFAEKNLNGVSGFIKQLISNGRQWSV 203
```

## RESULT 4

US-10-739-930-9287

; Sequence 9287, Application US/10739930

; Publication No. US20040216190A1

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH

; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT

; FILE REFERENCE: 38-21(53377)B

; CURRENT APPLICATION NUMBER: US/10/739,930

; CURRENT FILING DATE: 2003-12-18

; NUMBER OF SEQ ID NOS: 11088

; SEQ ID NO 9287

; LENGTH: 344

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: GLYMA-23APR03-C63260\_1.p

US-10-739-930-9287

```
Query Match 18.3%; Score 329; DB 17; Length 344;
Best Local Similarity 27.2%; Pred. No. 6.4e-22;
Matches 94; Conservative 57; Mismatches 130; Indels 64; Gaps 10;
```

```
Qy 6 FHKLIIFSSTIOEKRLRVPDKFVSKPKDLSVAVALTPDGHVWRVGLKADNKIWFQDGM 65
Db 16 FFKIITAHNVHEGKLMIPNKVKYKRLQNTLFLKTPNGAEWKMLKKRDGKIWFQKGM 75
```

```
Qy 66 QEFVDRYSIRIGYLLIFRYEGNSAFSVIFNLSHSEINHYSTGLMDSAHNHFKARLPED 125
Db 76 KEFAEYHSLAHGHLILFRWDVTSHFQVHIFDLSALEIEY----- 114
```

```
Qy 126 LEDEDAEVIFFSSVVPSPLESTVPANKGVASSAIOITLFTGPVKAEEPTTPKIPKKGR 185
Db 115 -----PTEI-----IKGTASNKGNESPGDEHLEC-----HRSGQ 145
```

```
Qy 186 KKNADPEEINSSAPRDD---DPEN-----RSKPYESASARK-RTVTAERERAINAAKT 236
Db 146 KRKNSVEFLQCCQMRGRKCVKVENTMILPQALHHTATKCKGSKAMDQNTALDRASS 205
```

```
Qy 237 FEPTNPFRRVLRPSYLYRGCMYLPSPGFAEKYLS--GISGFIKQV-LAEKQMPVRC-LY 292
Db 206 FKSCNPPFLTYMHRTHISSHGSNLPMKFCFCSHLDLHKRRLISLQVLSGRINWPAKYQIH 265
```

```
Qy 293 KAGRA-KFSQGWYFTELENNLGEQDVCVFELLRTDRFVLKVTAPR 336
```

```
Db 266 KQKTAIRFKLSWNAFVKDNNLKVGDVCFELVHGTKLTFVLVHIFR 310
```

## RESULT 5

```
US-10-425-115-274348
; Sequence 274348, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 274348
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(502)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: M174577_181795C.1.pap
US-10-425-115-274348

Query Match      16.1%; Score 289.5; DB 17; Length 502;
Best Local Similarity 26.2%; Pred. No. 5.2e-18;
Matches 106; Conservative 59; Mismatches 155; Indels 85; Gaps 17;

QY 4 PPFHKLIFSTIQERLRLVDPKFKDELVS-AVALTVPDGHVVRVGL--RKADNKIW 60
Db 98 PQFFKVFPEQSTE-RLRIPTWFOHLEQOPTGAVSLRGPNGRWQAALASESESAW 156
QY 61 -FDGQWQEVDRYSRIGVLLIFRYEGNSAFSVYIFNLHSHSE--INYHSTGLMDSAHNHF-----KRARLFEDLEDEDA 131
Db 157 CPDQGWKEFVTHSLRGLHFLVTRDGPQAFSVAVFS-----SSGVIDPAALDARP 207
QY 114 -HNHFKRARLFE-----DLEA--DEDAEV-IPSSVYVPSPLPESTVPANKGYASSAI 160
Db 208 TANGNAVXLEGEVGVGRGDVDAGDTSSEVSLIPAEEDGAGATGRTGATSGGAGNS 267
QY 161 QTLF-----TG-----PVKAEPTPTPKPKRGRKKKNA----- 190
Db 268 ENSVLVREGRGVTKRARATTSIDLPAASAFKKHSALAKKAGKRRPQAATSKDVSIMVH 327
QY 191 -----DPEINSAPRDDDPENRSKFYSAS-----AKRTVTAEERERAINAKT 236
Db 328 NPTFSLDESCTFNKTQIRDKNVPRSGKFLRTTKAPVVVISQRRRVTEEEKCLARKAME 387
QY 237 PPTPTPFRVRLPSVLYRGCIWYLPSPGFAEKVLSGISGFIKVQLA--AEKQWVRCVLYKAGRAKFSQ-GW 303
Db 388 FSKHPFTWQIMESVYVGFPMNACEFVRSLPRTSK--KWTLPDMPGKPDWDSVYXR 445
QY 294 AGRAKFSQ-GWYEFTLENNLGGDVCVFELLRTDFVLKVTAFRV 337
Db 446 RPRRLRLXGWRGALGNLXSKSDVCVFELFEDN--MRVHYRV 488

RESULT 6
US-10-425-114-63507
; Sequence 63507, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaka, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
```

```
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 63507
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73255A02_FLI.pap
US-10-425-114-63507

Query Match      15.6%; Score 280; DB 15; Length 375;
Best Local Similarity 27.7%; Pred. No. 2.7e-17;
Matches 89; Conservative 54; Mismatches 136; Indels 42; Gaps 13;

QY 19 RLRVPDKFVSKPKDELSVAVALTVP-DGHVWRVGLRKADNKIWFQDQWQEVDRYSIRIG 77
Db 59 RISVPEKFAKPKGQITTEFHLKSPSSAETWVGVGKGDGLFLVSGWENFAKAHLEEN 118
QY 78 VLLIFRYEGNSAFSVYIFNLHSHSE--INYHSTGLMDSAHNHF-----KRARLFEDLEDEDA 131
Db 119 DLLLFACSGNSSEFVLVFGASGCEKVSLSFGSLGPGDMGQFNDVVRHGVHHSVTVSDS 178
QY 132 EVIPFSSVPSPL-----PESTVPANKGYASSAIQTLPFGPVKAEPTPTPKPKRGRKK 188
Db 179 E----DTVAPSQLVRSPRNALPLKEPSGKAR-----PSKYESPSSNFIVRHVATGKE 227
QY 189 NADPEEINSAPRDDDPENRSKFYSASARKTVTAEERERAINAKTPEPTNPFRVVL 248
Db 228 GTDDEYANSN-----YYLSLSANR--LGDEEKEEIIIGLA-PIRPNPVFVTL 272
QY 249 RPSLYR--GCIWYLPSPGFAEKVLSGISGFIKVQL--AEKQWVRCVLYKAGRAKFSQ-GW 303
Db 273 RKNHVQRNNCLJ-IPSKFAADHLGERAHNIIILRRPNRKEKHVSVYTHSRHTCEQNLA 331
QY 304 YEFTLENNLGGDVCVFELLR 324
Db 332 PKFTRENKLHEGDIQVFEMLK 352

RESULT 7
US-10-437-963-157411
; Sequence 157411, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 157411
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_56986C.1.pap
US-10-437-963-157411

Query Match      15.5%; Score 278; DB 16; Length 402;
Best Local Similarity 26.8%; Pred. No. 4.6e-17;
Matches 103; Conservative 57; Mismatches 157; Indels 68; Gaps 16;

QY 3 RPFFHKLIFSSTIQERLRLVDPKFKV-----SKFDELSSVAVALTVPDGHVVRVGL 52
```

Db 18 RPFFKVLWGDF--KQRLKIPNFCFKHIPWBSRKAKGLKEASMAATLEGPSGRTWLVVI 75  
QY 53 RKADNKIFQDCWQGFVDRYSIRIGVLLIFRYEGNSAFSVYIFNLHSEINVHSTGLMDS 112  
Db 76 RTAGTFTTSWPKFVQDAULELEFVVFRVDGNTRFTAMVDFRTACE---REDLMGG 131  
QY 113 A---HNHFKRARLFEDLEDA-----EVI-----FSSVYP-----SPLPES 147  
Db 132 GGGDRPKRGRPRTAASRAARPKKDSVGKEMTYRASPGGQPLQIVDSWTPPEGS 191  
QY 148 TVPANKGYASSAIQILFTGPVKAEBPTTPKIPK-----KGRKKKNADPEEN- 196  
Db 192 TAVKNEEADAE---LPVCELPASSAPRHYPEGALDADGGAARGAATRSLODDLAL 247  
QY 197 SSAPRDDDPENSKFYESASAKRTVTAERERAINAAKTEPTNPFRRVLRPSLYRG 256  
Db 248 ASIP-----PSIRR--YKGVSRRAVATAERORATEIAHAFRSPDPCVIRMTWHVYS 301  
QY 257 CIMYLPSPGFAEYKYL-SGISGFIKVLQABKQWVRCLYKAG-RAKFSQGWYEFTELENNLGE 314  
Db 302 FMWRPTGFSRQHLPRRTDVLDPGGKWSV--LYIPNTEDRLSLRGWCAFAFGNCLEE 359  
QY 315 GNVCFVLELLRTRDFVLKVTAFRNE 339  
Db 360 GDYCVFELVAAAEF--RVHIFRWE 382

## RESULT 8

US-10-424-599-166413  
; Sequence 166413, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 166413  
; LENGTH: 181  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_121286C.1.pap  
US-10-424-599-166413

Query Match 14.8%; Score 266.5; DB 15; Length 181;  
Best Local Similarity 34.4%; Pred. No. 1.8e-16;  
Matches 64; Conservative 33; Mismatches 66; Indels 23; Gaps 5;  
QY 3 RPF-FHKLIFSSITQEKRLRPDKFVKFKDBLSVAVALTPDGHVWRVGLRKADNKIWF 61  
Db 13 KPFHFLITIAQLQDGLKLMINKFVKYEGGLPNALFLKTPNGTWNFNLEKHGDKLWF 72  
QY 62 QDQWQFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLHSEINVHSTGLMDSA---HNHF 118  
Db 73 QKGWKEFAEYHSLAHGLLVFRHGTSHFQVHIFDLSSLEIDYPSKGTGKTSNHEGNK 132  
QY 119 RARLFEDLEDAEVIFFSSVYPSLPSTVTPKNGYASSAIQTLTGPVKAEBPTTPK 178  
Db 133 QPR-----NEENLEYLQPYQV-----RSHSKSVKVENMMTLPEEAQPTDTFK 174  
QY 179 IPKRG 184  
Db 175 -EKS 179

## RESULT 9

US-10-437-963-192114  
; Sequence 192114, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 192114  
; LENGTH: 362  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_8836C.1.pap  
US-10-437-963-192114

; Sequence 192114, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 192114  
; LENGTH: 362  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_8836C.1.pap  
US-10-437-963-192114

Query Match 14.8%; Score 265; DB 16; Length 362;  
Best Local Similarity 26.6%; Pred. No. 6.4e-16;  
Matches 97; Conservative 57; Mismatches 148; Indels 62; Gaps 13;  
QY 20 LRPDPKFKVSKFKDBLSVAVALTPDGHVWRVGLRKADNKIWFQDQWQFVDRYSIRIGYL 79  
Db 1 MKIPSSFNQCLQNPQTMVSLVDRSGNKSABELTSDSGFPFVHGKWEFVRDINSIQCGQF 60  
QY 80 LIFRYEGNSAFSVYIFNLHSEINVHSTGLMDSAHNHFKRARLFEDLEDAEVIFFSSV 139  
Db 61 LVFTYDKRSQFSVTF--EPSGIDKIST---FSAHP--SKNVIKTESDEGMVTAATT 113  
QY 140 YPSLPPESTVPANK-----GYASSAIQTLFTGPV---KAEETPTPKPKR-----G 184  
Db 114 -----EKMAPALKENNGITGKRTDVLDMEDRVVFKSSEANVCSSRRKAGASAG 167  
QY 185 RKK-----KNADPEEINSAPRDDDPENRKFYESASAR----- 218  
Db 168 KSKVTSTSHNSTRGSCSSDEDNSS--KSPNPPFLMRFLSGEVSRRGRCSVGQRLTVI 226  
QY 219 --KRTVTAERERAINAAKTEPTNPFRRVLRPSLYRGCMYLPSPGFAEYKLSIGSF 276  
Db 227 SQRRPVTEAEKDHALQAREFKSKNPPFAVQIMMSYVYVGFPMNIPCFVRECLPHTNKR 286  
QY 277 IKVOLAE-KOMPVRCLYKAGRA--KFSQGWYEFTELENNLGEQDVCVFELLRTRDFVLKVT 333  
Db 287 ITLWDPOKAEVNVYVYSDRSVSGSGGWKPAVGNLKEKFDVCVFELVQKDN--IKVH 344  
QY 334 AFRV 337  
Db 345 IYRV 348

RESULT 10  
US-10-437-963-178183  
; Sequence 178183, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 178183  
; LENGTH: 362  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_8836C.1.pap  
US-10-437-963-178183

```
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 178183
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_75764C.1.pap
US-10-437-963-178183

Query Match          14.0%; Score 251.5; DB 16; Length 750;
Best Local Similarity 26.9%; Pred. No. 3.2e-14;
Matches 86; Conservative 55; Mismatches 134; Indels 45; Gaps 13;

QY 22 VPKDFVSKFDELVAVALTVPDGHVVRVGLAKAD-NKILWFODGQWQEFVDRYSIRIGYLL 80
DB 424 VFARFANNFNHISEVNLRSGETWSIGVANSAGELVLPQWKEFVDGNGIEGDCL 483
QY 81 IFRYEG-NSAFSVYIFNLSHSEINYHSTGLMDSAHNHFKARLFEDLEDE-DAEVIFFPS 138
DB 484 LFRYGVSSSPDLIFDPSCGE-----KASPHFVSGHGFGRASAGAEQGGNG 533
QY 139 VYPSLPPESTVPANKGYASSAIQTLFTGPVKAEBEPTTPKIPK--KRG-----RKKN 189
DB 534 RRTPII-----VGDNGHRRHLEMTLHRNSCRS-----IPRACKRSIFSDETEAKEND 581
QY 190 ADPEINSSAPRDDDPENRSKFVESASAR--KRTVTAERERAINAAKTFFPTNPFRV 247
DB 582 GEDEVDVAAA-----EGRGYGETYFNRHGRVAEYNLEEDREISRPVVPQGNPVQV 637
QY 248 LRPSYL--YRGCIMYLPSPGAFKYLKSGISGFIKVLAER--QWPVRCLYKAG-RAKFSQG 302
DB 638 IHSSHVRSSKYCIVGSPPEFAGKYLCAVEREVLERASRGGEHVPFVHRQNRGFGYAG 697
QY 303 WYEFLENNLGGDVCVFLP 322
DB 698 WRQFAGDNRNLVAHVDCLPEL 717

RESULT 11
US-10-437-963-203869
; Sequence 203869, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 203869
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_99011C.1.pap
US-10-437-963-203869

Query Match          13.7%; Score 246; DB 16; Length 737;
Best Local Similarity 23.8%; Pred. No. 1e-13;
Matches 74; Conservative 60; Mismatches 123; Indels 54; Gaps 8;

QY 5 FFHKLIFSSTIQEKLRLVPDKFVSKFDELVAVALTVPDGHVVRVGLRKADNKIWFQDG 64
```

```
DB 149 YFFKVMIGGF--RRQMTIPYFAENFRDQIQGTIKLKAENGNTCSVLVDKCNKLVLTGK 206
QY 65 WQEFVDRYSIRIGYLLIFRYEGNSAFSVYIFNLSHSEINYHSTGLMDSA-HNHFKARLUF 123
DB 207 WAEFANSHDIKMGDFLVFRYTGNSQFEVKIFDPS-----GCVKAASHNAVIGQHA 257
QY 124 EDLEDEDAEVIFFPSVYPSLPPESTVPANKGYASSAIQTLFTGPVKAEBEPTTPKIPK 183
DB 258 QNMQGDPIEILSCSDHLRAQSLTTERQNO-----PEKD 291
QY 184 GRKKQVADPEINSSAPRDDDPENRSKFVESASARKRTVTAERERAINAAKTFFPTNP 243
DB 292 VIDNCKNKMKTEHASSSEDDQ-----ETPTAEVHRMKVEMVRAIHS-----NHFP 337
QY 244 FRVLRLPSLYR-GCIMYLPSPGAFKYLKSGISGFIKVLAERKQWVPR-CLYKAGRAKFSQ 301
DB 338 FVAVWKKSNVTPQPCYVAISRYANEYFPGDQMLTLQRHGRWQVKFCISRKRLMLSK 397
QY 302 GWYEFLENNL 312
DB 398 GWRKFTRDNEL 408

RESULT 12
US-10-437-963-152254
; Sequence 152254, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 152254
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_52323C.1.pap
US-10-437-963-152254

Query Match          13.4%; Score 241.5; DB 16; Length 322;
Best Local Similarity 25.2%; Pred. No. 8.5e-14;
Matches 80; Conservative 57; Mismatches 140; Indels 41; Gaps 8;

QY 18 KRLRPDPFVSKFDELVAVALTVPDGHVVRVGLRKADNKIWFQDGQWQEFVDRYSIRIG 77
DB 37 ERMIIPNEFLQYFRGKIPRTIKLQRDCTVDVQVTKNLKISLQSGMKAFVTAHDLQNG 96
QY 78 YLLIFRYEGNSAFSVYIFNLSHSEINYHSTGLMDSAHNHFKARLFEDEDAEVIFFPS 137
DB 97 DFLVPSYDGIKSLKVLIFGPSGCEKVSRSRSLKNATHCGEK----- 137
QY 138 SVYPSLPPESTVPANKGYASSAIQTLFTGPVKAEBEPTTPKIPKGRKKKNADPEINS 197
DB 138 --WEEPL-----HISNSHDL-----PVKS---PQVSKSEKQWDSSEQNDANTIEE 180
QY 198 SAPRDDDPENRSKFVESASARKT-VTAERERAINAAKTFFPTNPFRVLRPSLY-R 255
DB 181 VALQGDLLQGHVPL--NCILPKHTLTDQKQLESKVGAHSEIPIYGCILRKSRVHGK 238
QY 256 GCIMYLPSPGAFKYLKSGISGFIKVLAERKQWVPRCLYKAGRAK-FSQGWYEFTLENNLGE 314
```

```
Db 239 SQTVDICREYADVLPFKELNMTLQRHKGKWEVLCTKDTKRLSTGWSRPAQENNLQV 298
QY 315 GDVCFPELLRTRDFVLKV 332
Db 299 GDICLFELLKKKEYSMNV 316

RESULT 13
US-10-437-963-129669
; Sequence 129669, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 129669
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_31904C.1.pap
US-10-437-963-129669

Query Match 11.9%; Score 214; DB 16; Length 306;
Best Local Similarity 19.7%; Pred. No. 2.9e-11;
Matches 63; Conservative 55; Mismatches 137; Indels 64; Gaps 5;
QY 22 VPDKFVSKFDELVAVALTPDGHVWRVGLKADNKIWFQDQWQEFVDRYSIRIGYLLI 81
Db 46 IPNEFLHNFGGKIPKSIKLETRSGTLFDVQVTKNSGRVQLQSGWASYSAHDLKIGDFLV 105
QY 82 FRYEGNSAFSVIFNLSHSEINHYSTGLMSAHNFKRRLFEDLEDEDAEVIFFSSVYP 141
Db 106 FKYSGSQKTLIFD----- 120
QY 142 SPLPESTVPANKGYASSA1QTLTGPVKAEEPTTPPKPKKGRKKQNADPEINSSAPR 201
Db 121 -----SSGCEKVCKEPDMGSRSDYDIAMENSQDEKKKQKQDISRQGTVK 165
QY 202 DDDPENRSKFYESASARKET-VTAERERAINAAKTFEPTNPFERVLRLPSLYLR-GCIM 259
Db 166 PSEGLKAELVPGCILPSRDLTLQKNILIEKVAINSETPIYGVVMNSSIHGIPCTV 225
QY 260 YLPSGPAEKYLSGISGFIKVLAEKQWVR-CLYKAGRAKFSQGWYFETLNNLGEQDVC 318
Db 226 EISKYADVLPEDGTVLVQLHKGKSMVRCCCLTKQNSKRFKQWQFQAGDNKLHLGDI 285
QY 319 VFELLR-TRDFVLKVTAFR 336
Db 286 LFDLLKDKKKYVMDVHIIR 304

RESULT 14
US-10-424-599-206969
; Sequence 206969, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
```

```
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 206969
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_2891C.1.pap
US-10-424-599-206969

Query Match 11.7%; Score 210.5; DB 15; Length 164;
Best Local Similarity 42.1%; Pred. No. 2.6e-11;
Matches 45; Conservative 18; Mismatches 43; Indels 1; Gaps 1;
QY 2 PRPFHKLIFSSTIQEKRLVRPDKFVSKFDELVAVALTPDGHVWRVGLKADNKIWF 61
Db 57 PAVHFVKIILTSLADGIL-LPKKFKTKYGDGMSNPVFLKPADGTGTEWKIHYTKHGSEIWF 115
QY 62 QDGHQEFVDRYSIRIGYLLIFRYEGNSAFSVIFNLSHSEINHYSTG 108
Db 116 QKGWKEFATYYSLDHGHLLFFEYEGTSHFDVHLFDSALEIDIPSHG 162

RESULT 15
US-10-425-115-314879
; Sequence 314879, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 314879
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(396)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_50230C.1.pap
US-10-425-115-314879

Query Match 11.4%; Score 204.5; DB 17; Length 396;
Best Local Similarity 22.2%; Pred. No. 3.2e-10;
Matches 76; Conservative 52; Mismatches 104; Indels 11; Gaps 13;
QY 3 RPFFHKLIFSSTIQEKRLVRPDKFVSKFQDE--LSVAVALTV-PDGHVWRVGLKADNKI 59
Db 136 KPQFIR-IFHNLSEK-WMLPARVKQVSEARLNSRIAIIFSPIAKFWRVELKNDQSGI 193
QY 60 WPDQWQEFVDRYSIRIGYLLIFRYEGNSAFSVIFNLSHSEINHYSTGLMSAHNFKR 119
Db 194 FFTGWSQFLDPHGKNGEVLLKYEWNVFKFAFLSGCQ----- 235
QY 120 ARLPELDEDEDAEVIFFSSVYPSPLPESTVPANKGYASSA1QTLTGPVKAEEPTPKI 179
Db 236 ----KDFKNQAGI-----QINTKKQOETPPI 258
QY 180 PKKGRKKKNADPEE-----INSSAPRDDDDPENRSKFYESASARKETVT-AEERAIN 232
Db 259 RKRKSNDEKSSSEENKRLKTPVTPSPSPDPSLEKPYQIGTSSWIKKRIINTYVALEQFLALS 318
```

Qy	233	AAKTFETNPFRVVLPSYLYRGCMVLPSPGFAEKYLSGISGF	-----IKVQLA-----EK	284
			: : :	
Db	319	V-----KFCNSI-GFEMACTITLTKTATNSTR	343	
			: : :	
Qy	285	QWPVR-CLYKAGRAKFSQGWYFTELENNLGGSDVCVPFELLATR	326	
			: : :	
Db	344	SWQVRGTAYKWHRYIIIGVGWKSFCQDNRLLKAGDLCAFNIIEXR	386	
			: : :	

Search completed: December 30, 2004, 06:22:52  
Job time : 368 secs

**THIS PAGE BLANK (USPTO)**